Journal of
CHRONIC
DISEASES

VOLUME 10 . NUMBER 3

SEPTEMBER, 1959

OL.
10
959

## Cardiovascular

# VENOUS PRESSURE, EMOTIONS, AND CONGESTIVE HEART FAILURE

KERR L. WHITE, M.D.,
DAN A. MARTIN, M.D.,* AND
CHARLES R. VERNON, M.D.,
CHAPEL HILL, N. C.

*From the Departments of Medicine, Preventive Medicine, and Psychiatry, School of Medicine, University of North Carolina*

(Received for publication March 22, 1959)

## INTRODUCTION

ALTHOUGH much is known about overt congestive heart failure, less is known about the physiologic, psychologic, and environmental circumstances associated with its initiation. In 104 episodes of congestive failure which occurred in 100 consecutive patients, Sodeman and Burch[1] were able to identify precipitating factors in 55 patients, and, in 40, cardiac decompensation was treated effectively and rapidly (Fig. 1). In contrast, of the 47 patients in whom no precipitating factors were demonstrated, 46, because of their persistent congestive failure, were considered to have a poor prognosis. "Psychic trauma" was considered to have precipitated one episode of failure in each of 2 patients. A series of 1,000 patients with congestive heart failure was reviewed by Boyer, Leach, and White.[2] In 561 of these cases, precipitating factors were identified but in the remainder no such factors were recorded. "Emotions" were believed to have initiated failure in one patient. In almost half of the patients in these two series there were no apparent explanations for the occurrence of heart failure, either in the many patients with an insidious onset or in those with acute onset.

Supported by research grants from the National Heart Institute (H-2360), United States Public Health Service, and from the Durham-Orange County Heart Association and the United Medical Research Foundation of North Carolina.

*This work was conducted during the tenures of Research Fellowships from the American College of Physicians and the American Heart Association.

This paper was presented in part before the Council on Circulation of the American Heart Association at the Thirty-First Scientific Sessions, San Francisco, Oct. 24, 1958.

In neither of the preceding two series was any special effort made to ascertain whether less apparent emotional factors may have been associated with the onset of cardiac failure. Chambers and Reiser,[3] who were particularly interested in correlating emotional factors with the onset or exacerbation of congestive heart failure, studied 25 consecutive patients admitted to a general hospital with this diagnosis. In 23 patients, precipitating factors were identified, and in 19 "it was felt that emotional tension was the factor immediately responsible for increasing the workload beyond cardiac capacity." The marked discrepancies between the identification of 56 per cent of the precipitating factors (614 out of 1,100 patients) in the first two series, and of 92 per cent (23 out of 25 patients) in the third series, and between the reported incidence of emotional disturbances associated with cardiac failure of 0.27 per cent and 76 per cent, respectively, emphasize the need for further study of these problems.

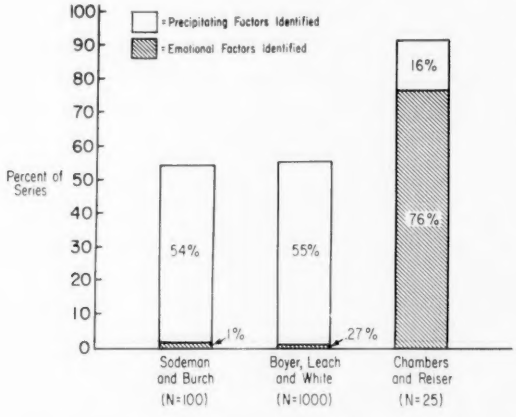


Fig. 1.—Three series of patients with congestive heart failure in which attempts were made to determine the precipitating factors. There is a marked discrepancy between the incidence of emotional factors in the first two series of Sodeman and Burch[1] and Boyer, Leach, and White[2] and the third series of Chambers and Reiser.[3]

There may be several mechanisms by which emotional influences participate in the development of congestive failure. Arrhythmias, including auricular and ventricular ectopic beats,[4] paroxysmal auricular tachycardia and fibrillation,[5] and paroxysmal ventricular tachycardia,[6,7] have been observed in association with emotionally disturbing situations. The tachycardia accompanying anxiety in patients with mitral stenosis may result in decreased diastolic filling, a diminution in cardiac output, and an increase in pulmonary engorgement with edema and dyspnea.

Grollman[8] observed that normal volunteers subjected to potentially threatening situations responded with varying increases in cardiac output as measured by an indirect Fick method. He concluded that cardiac output was increased in proportion to the "degree of anxiety" reported by each subject. Hickam, Cargill, and Golden,[9] using a ballistocardiograph, studied medical students and demonstrated changes in cardiac output in association with "anxiety" preceding

an examination. The direction of the changes in cardiac output varied, but most subjects showed an increase. Three patients in cardiac failure also were studied and their cardiac functions compared during "anxiety" and "composure." Oxygen consumption increased slightly during the period of anxiety but cardiac output, as determined by the direct Fick method, did not change or decreased slightly. Pulmonary arterial pressure increased by 50 per cent in one patient during the period of anxiety. Stead, Warren, Merrill, and Brannon[10] measured right atrial pressure and cardiac output during 7 experiments in 5 normal subjects, all of whom had pulse rates greater than 82 per minute and basal metabolic rates greater than "plus 10." The latter two variables were considered evidence of "anxiety." Both cardiac output and oxygen consumption increased in all subjects, the former being out of proportion to the latter, and occurring without measurable change in right atrial pressure. Stevenson, Duncan, and Wolff[11] observed that individuals with normal, and those with abnormal, hearts recovered from exercise less efficiently during states of emotional tension than when relaxed. Apart from other studies of peripheral arterial pressure, little else is known about the effects of emotional disturbances on cardiac function.

On the venous side, "apprehension"[12,13] and certain "mental states" including "thought," "interest," "anger," "fear," and "dozing" have been reported to affect venous tone.[14] "Fear" was sometimes associated with an increase and sometimes with a decrease, "dozing" with a decrease, and all others with increases in tone. That these changes are probably mediated through the sympathetic nervous system's control of both arterial and venous reactivity is supported by the work of Duggan and associates,[12] Salzman,[15] and Alexander.[16]

In an attempt to understand better the frequently unexplained onset of congestive heart failure, a series of studies was undertaken. Specifically, they were designed to (1) examine the effects of emotional stimuli on (a) central venous pressure and (b) peripheral venous tone, and (2) evaluate the significance of specific emotional experiences in relation to the symptoms of congestive heart failure.

## METHOD

Central venous pressure (CVP) was measured by means of a calibrated polyethylene catheter (Clay-Adams No. PE50) filled with heparin (10 mg. per cubic centimeter) inserted into an antecubital vein through a No. 17 needle under local 1 per cent procaine anesthesia and advanced to the central veins. The needle was removed from the vein over the catheter and the latter connected to a 3-way stopcock attached to a Statham P-23 BB 0-5 cm. Hg strain gauge. The system was filled with a dilute solution of heparin in isotonic saline (2 mg. per cubic centimeter) (which is said not to affect venous reactivity[13]) and flushed out at appropriate intervals. The position of the catheter tip was estimated by external measurement and confirmed by the configuration of the pressure curve following a Valsalva maneuver.[13] The strain gauge was placed at the level of the right auricle and the system standardized by means of a Burch-Winsor phlebomanometer.[17] Peripheral venous segment pressure (PVSP) was measured similarly, except that the catheter tip remained in the forearm vein. When venous pressure

within an isolated segment was measured, the limits of a segment without "feeders" were determined by alternately filling and emptying the vein below a recognizable valve and visually establishing its competence.[18] Two specially designed Plexiglas occluders of different sizes (Fig. 2) applied a constant pressure to either end of the venous segment with minimal pressure on the soft tissues of the forearm. Discomfort was negligible for periods up to 30 minutes. Persistence of a high steady pressure in the distended isolated segment confirmed its isolation.

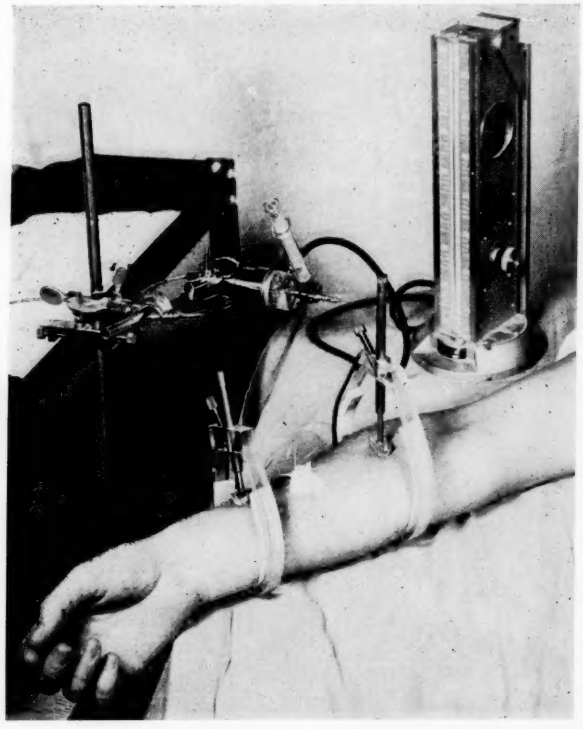

Fig. 2.—Plexiglas occluders in place with polytehylene catheter attached to strain gauge. The Burch-Winsor[17] phlebomanometer used to standardize the system is shown in the background.

Respiratory rate and pattern were measured by means of a Sanborn Model 108 pneumograph connected to a Statham P-23 BB 0-5 cm. Hg strain gauge. Cardiac rate and rhythm were recorded from Lead II of the electrocardiogram. In most patients two electrodes attached to the right leg recorded the psychogalvanic skin response, a measure of autonomic excitability believed to be independent of vasomotor activity but under hypothalamic control.[19] An event marker indicated the onset and cessation of various stimuli.

All instrumentation was purposely kept simple in order not to encumber the patient unduly and was attached expeditiously as the patient rested quietly on a hospital bed in an air-conditioned room. For dyspneic patients, the head of

the bed was elevated and the height of the strain gauge adjusted visually to the level of the right auricle. The elevation of bed and strain gauge was measured so that similar conditions would prevail on successive studies of the same patient.

In an adjacent room from which the patient was observed, data were monitored and recorded continuously on an 8-channel Offner D-3 electroencephalograph run at a paper speed of 1 or 2.5 mm. per second. An adequate base line was recorded, although the relative constancy of the base line does not necessarily imply that the patient was free from anxiety,[20] nor does it provide any data about the absolute tone of the veins.

## OBSERVATIONS

Two related but different sets of experiments were conducted in an attempt to ascertain the relationships between emotional stimuli and venous pressure changes. One set was directed toward changes in CVP and the other to changes in PVSP and tone.

### CENTRAL VENOUS PRESSURE OBSERVATIONS

*Patients.*—Twenty-five studies of CVP changes were conducted according to the techniques described above. Twenty-two patients were studied, 3 being studied twice to verify the reproducibility of the method. The series included patients with arteriosclerotic heart disease, hypertensive cardiovascular disease, rheumatic heart disease, idiopathic and tuberculous pericarditis, chronic cor pulmonale, congenital heart disease, and 5 with apparently normal cardiovascular systems. Table I compares the base-line CVP of the 25 studies with the severity of heart disease, classified in accordance with the New York Heart Association's criteria. The "normal" value for this method is about 60 mm. of water. The patients with more seriously decompensated hearts (Class IV) with higher base-line pressures often required elevation of the head of the bed, and appropriate adjustment of the strain gauge level.

*Physical Stimuli.*—In order to assess variations in CVP associated with different physical stimuli, and as a basis for the ready identification of these influences and for their comparison with responses to emotional stimuli, each patient performed the following activities: Valsalva maneuver, maximal inspiration and expiration, counting aloud, reading aloud, lifting a 2-Kg. weight 15 times at arm's length with the left hand, rapid flexion and extension of the feet, elevation of the left arm, turning of the head to the left and then to the right, and elevation of the head on the neck.

*Examples:* Fig. 3 illustrates in a patient with hypertensive cardiovascular disease, Class III, the effects on CVP, heart rate, ECG, and respiratory pattern of: (*A*) reading aloud a "neutral" paragraph from a newspaper, (*B*) counting aloud, (*C*) the customary response to the Valsalva maneuver showing an immediate increase in CVP and a slight decrease in heart rate, and (*D*) the response to an emotionally charged rhetorical question.

*Comment:* In general, it may be said that all maneuvers produced changes in CVP which could be identified either by the contour of the pressure curve, by concomitant changes in rate or patterns of respiration, by changes in heart rate,

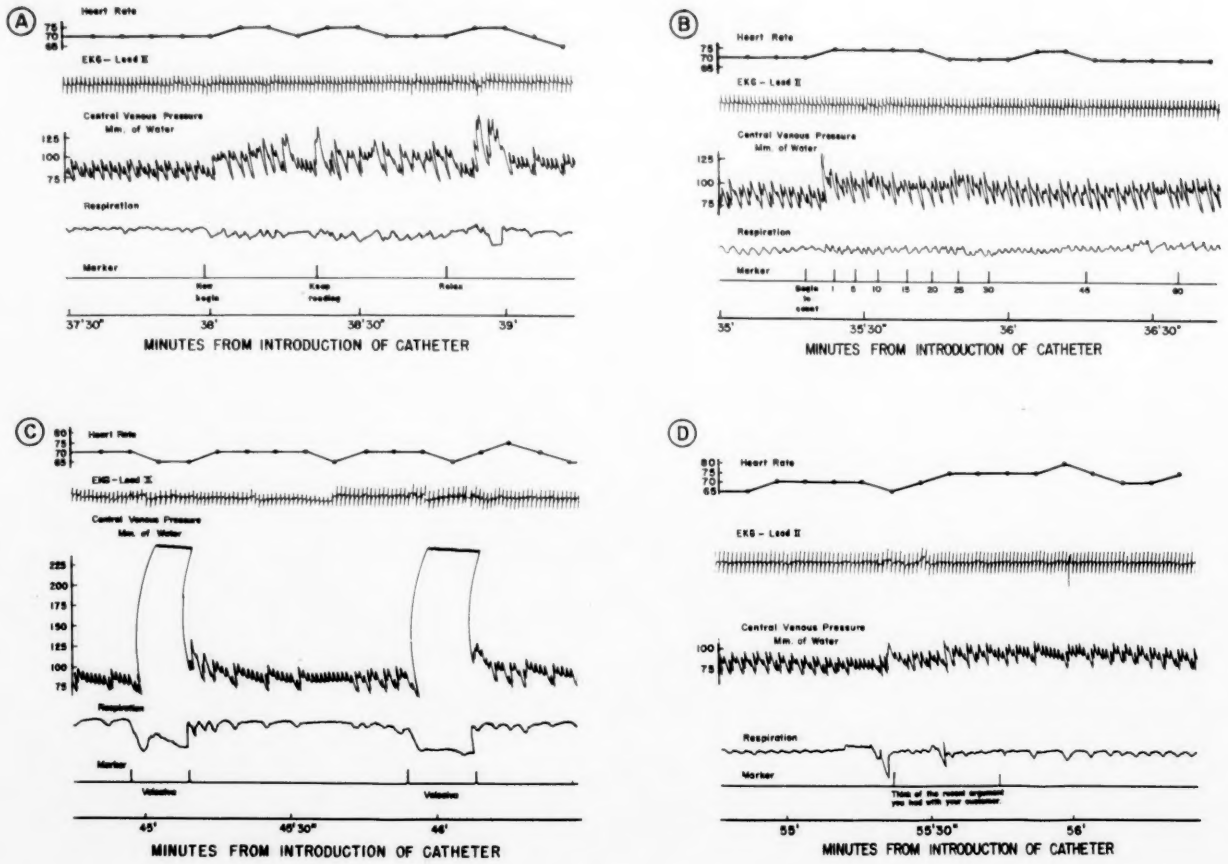


Fig. 3.—Four tracings from a patient moderately ill with hypertensive cardiovascular disease, Class III, showing the effects on heart rate, ECG-Lead II, central venous pressure (CVP), and respiratory pattern of three physical stimuli: (*A*) reading aloud, (*B*) counting aloud, and (*C*) the Valsalva maneuver, and (*D*) of one emotionally disturbing stimulus: a question about a recent argument.

VOL. 10 1959

TABLE I. BASE-LINE CENTRAL VENOUS PRESSURE (CVP) DETERMINATIONS (MILLIMETERS OF WATER) IN 25 STUDIES OF 22 PATIENTS

| TYPE OF HEART DISEASE | PATIENT NUMBER | SEVERITY OF HEART DISEASE* | | | | |
|---|---|---|---|---|---|---|
| | | NORMAL | I | II | III | IV |
| Normal (no heart disease) | 15 | 54 | | | | |
| | 5 | 60 | | | | |
| | 7 | 60 | | | | |
| | 11 | 70 | | | | |
| | 22 | 92 | | | | |
| Hypertensive | 13 | | | 34 | | |
| | 16 | | | 90 | | |
| Rheumatic | 17 | | 35 | | | |
| | 2 | | 45 | | | |
| | 19 | | 85 | | | |
| | 20 | | | | 40[a] | |
| | 20 | | | | 40[a] | |
| | 12 | | | | 60[b] | |
| | 12 | | | | 65[b] | |
| | 4 | | | | 122 | |
| | 3 | | | | | 138 |
| | 6 | | | | | 145 |
| Arteriosclerotic | 21 | | | | 43 | |
| | 18 | | | | 125 | |
| | 10 | | | | | 100[c] |
| | 10 | | | | | 100[c] |
| Pericarditis | 8 | | | 115 | | |
| | 1 | | | 136 | | |
| Chronic cor pulmonale | 9 | | | 220 | | |
| Congenital heart disease (IVSD) | 14 | | 0 | | | |

*New York Heart Association Classification.
a, b, c, Two separate studies in the same patient.

or by evidences of muscle activity recorded in Lead II of the electrocardiogram. The observation of responses to physical stimuli provided typical patterns which served as comparisons for evaluating the emotional stimuli and which distinguished the latter from artifacts. In some instances, both physical and emotional stimuli appeared to contribute to the total venous pressure responses.

*Emotional Stimuli.*—Three types of observations were made to evaluate the effects of emotional stimuli on CVP and their possible role in the onset or exacerbation of congestive heart failure in patients with various forms of heart disease.

(1) *Predictive studies:* The hypotheses underlying this group of studies state: (a) that during the course of clinical history taking, emotionally charged experiences of patients can be identified readily; (b) that those experiences which are considered by the physician to be emotionally disturbing for each specific individual patient (referred to henceforth as "stressful") can be predictively distinguished from those experiences not emotionally disturbing for the same individual (referred to henceforth as "neutral"); (c) that the experiences pre-

dicted to be "stressful" will be associated with increases in CVP more frequently than those predicted to be "neutral"; and (d) that the predicted increases in CVP will be greatest in the patients with the most severe cardiac decompensation.

One of the investigators (D.A.M.) interviewed each of 15 patients briefly to determine some of the recent and remote life situations they had experienced. Based on this material, rhetorical questions were formulated which it was predicted would represent either "neutral" or "stressful" emotional stimuli for the particular patient. Upon completion of the studies employing physical stimuli in the laboratory and after a five-minute rest period, the patient was instructed not to speak but to listen carefully and to think about the rhetorical questions he would be asked. Before and after each question the base line was allowed to become stabilized, although it did not always return to the level immediately preceding each question within a reasonable period of time (i.e., 2-5 minutes). The questions were asked over an intercommunication system, the "neutral" questions usually being asked first, followed by the "stressful" questions. In several patients the sequence was varied and questions previously associated with an increase in CVP were repeated toward the end of the study. The second responses usually duplicated the initial ones. The records were analyzed without knowledge of the predicted "stressful" or "neutral" quality of each question, so that potential bias in deciding whether a specific CVP response was "positive" was minimized.

VOL. 10 1959

A readily apparent increase in mean CVP of any magnitude (range 5-15 mm. of water) following a question was considered a "positive" response. In this series of studies moderate changes in respiratory rate and pattern in association with an increase in CVP were acceptable, but a cough, arm movement, or muscle tension as detected from the respiratory tracing or the ECG excluded a response from being considered "positive."

*Example:* Fig. 4 shows the response of a 35-year-old poverty-stricken millworker unable to hold a job and harassed by a large family. He had rheumatic heart disease with mitral stenosis and his heart was well compensated clinically (Class I). In response to a question relating to the loss of his job and predicted to be "stressful" he showed perception of the question by a drop in the psychogalvanic skin response and experienced a modest but definite increase (10 mm. of water) in CVP lasting 35 seconds without any apparent change in heart rate or respiratory activity.

*Comment:* This "positive" response of relatively short duration in a patient with well-compensated heart disease suggests that he reacted to the "stressful" stimulus with little more than a transient increase in CVP. He showed more cardiovascular responsiveness than patients with normal hearts and had CVP increases in response to two out of three "stressful" questions and to neither of two "neutral" questions.

*Example:* Fig. 3, *D* shows the response to a question predicted to be "stressful" in the hypertensive patient (Class III) discussed above. The question relates to an argument leading to a fight which appeared to be temporally associated with exacerbation of the patient's congestive failure. The increase in CVP of 15 mm. of water persisted for 2½ minutes and was considered a

"positive" response. An initial deep inspiration immediately precedes the question and there is a paradoxical response in which the customary decrease in CVP associated with inspiration is replaced by an increase.

*Comment:* This "positive" response persists for longer than those seen in patients with better cardiac compensation, and continues after stabilization of the respiratory pattern. The duration of the response may be related either to the intensity of the emotional stimulus, the degree of cardiac decompensation, or a combination of both. Table II and Fig. 5 summarize the results of 16 studies in which the rhetorical "stressful" and "neutral" questions were employed in 15 patients. In 3 cases the patients did not respond to any question, either "neutral"

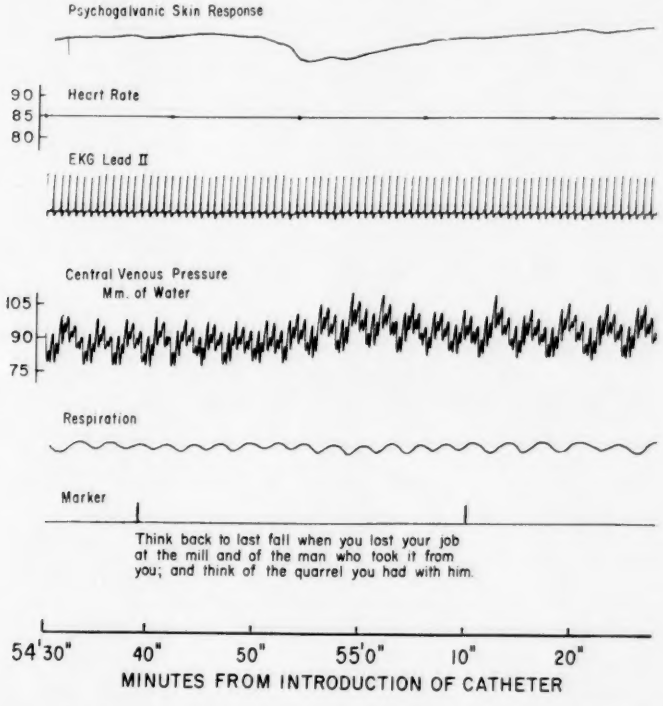


Fig. 4.—Psychogalvanic skin response, heart rate, ECG-Lead II, central venous pressure (CVP), and respiration in a patient with well-compensated rheumatic heart disease, mitral stenosis, Class I. There is a definite decrease in psychogalvanic skin response and a modest increase in CVP, but no apparent change in the other variables in response to an emotionally "stressful" question.

or "stressful," indicating either that all "stressful" questions were erroneously predicted, that the patients did not understand the instructions given them, that they did not understand the questions, or that they did not perceive them intellectually or react to them emotionally. One patient responded to all questions, both "neutral" and "stressful," suggesting the presence of considerable anxiety in relation to the experimental situation or the experience of being asked questions of any type. In 11 experiments the patients responded differentially

to "neutral" and "stressful" questions. There were 10 "positive" responses to 31 "neutral" questions and 39 "positive" responses to 70 "stressful" questions. Deviations from the hypothesis of independence as large as these or larger could occur less than 5 per cent of the time as determined by a chi-square contingency test.

TABLE II. INCREASES IN CENTRAL VENOUS PRESSURE IN RESPONSE TO RHETORICAL QUESTIONS

| PATIENTS | "NEUTRAL" QUESTIONS* | | | "STRESSFUL" QUESTIONS† | | | TOTALS | | |
|---|---|---|---|---|---|---|---|---|---|
| | NUMBER | RESPONSES | RATIO | NUMBER | RESPONSES | RATIO | NUMBER | RESPONSES | RATIO |
| *Nonresponders* | | | | | | | | | |
| J.T. | 2 | 0 | 0.00 | 3 | 0 | 0.00 | 5 | 0 | 0.00 |
| H.O. | 1 | 0 | 0.00 | 4 | 0 | 0.00 | 5 | 0 | 0.00 |
| B.T. | 3 | 0 | 0.00 | 10 | 0 | 0.00 | 13 | 0 | 0.00 |
| Total | 6 | 0 | 0.00 | 17 | 0 | 0.00 | 23 | 0 | 0.00 |
| *Total Responders* | | | | | | | | | |
| I.K. | 2 | 2 | 1.00 | 6 | 6 | 1.00 | 8 | 8 | 1.00 |
| *Differential Responders* | | | | | | | | | |
| L.H. | 3 | 0 | 0.00 | 6 | 1 | 0.17 | 9 | 1 | 0.11 |
| E.J. | 3 | 3 | 1.00 | 8 | 5 | 0.63 | 11 | 8 | 0.73 |
| E.B. | 2 | 0 | 0.00 | 3 | 2 | 0.67 | 5 | 2 | 0.40 |
| K.H. | 4 | 0 | 0.00 | 7 | 3 | 0.43 | 11 | 3 | 0.27 |
| C.M. | 2 | 0 | 0.00 | 5 | 2 | 0.40 | 7 | 2 | 0.29 |
| M.T. | 3 | 2 | 0.67 | 8 | 7 | 0.88 | 11 | 9 | 0.82 |
| N.E. | 4 | 1 | 0.25 | 10 | 5 | 0.50 | 14 | 6 | 0.43 |
| N.G. | 3 | 2 | 0.67 | 8 | 6 | 0.75 | 11 | 8 | 0.73 |
| M.W. | 2 | 0 | 0.00 | 3 | 1 | 0.33 | 5 | 1 | 0.20 |
| B.T. | 2 | 0 | 0.00 | 5 | 2 | 0.40 | 7 | 2 | 0.29 |
| N.B. | 3 | 2 | 0.67 | 7 | 5 | 0.71 | 10 | 7 | 0.70 |
| Total‡ | 31 | 10 | 0.32 | 70 | 39 | 0.56 | 101 | 49 | 0.49 |

*Responses to rhetorical questions which it was predicted would be "neutral" and to which the patient *would not* respond with increases in central venous pressure.

†Responses to rhetorical questions which it was predicted would be "stressful" and to which the patient *would* respond with increases in central venous pressure.

‡Number of responses to "neutral" and "stressful" questions in differential responders:

| | NEUTRAL | STRESSFUL | TOTAL |
|---|---|---|---|
| Responses | 10 | 39 | 49 |
| No responses | 21 | 31 | 52 |
| Total | 31 | 70 | 101 |
| | $\chi^2 = 4.732$ | $P < .05$ | |

Deviations from the hypothesis of independence as large as those observed or larger could occur less than 5 per cent of the time as determined by a chi-square contingency test.

The data from these studies appear to be compatible with the following conclusions: (a) Experiences which a patient with heart disease feels to be emotionally "stressful" can be elicited and predicted with a fair degree of confidence during a short medical interview. (b) In patients with varying degrees of congestive heart failure, rhetorical questions which the patient considers intellectually and emotionally, without speaking, are associated with varying but definite

increases in central venous pressure. Questions predicted to be "stressful" are more frequently associated with increases in CVP than questions predicted to be "neutral."

Alterations in respiratory rate, depth, and rhythm, and in cardiac rate and rhythm, may be among the mechanisms contributing to the observed increases in CVP but do not appear to be involved in every instance, and there is some evidence that CVP increases before heart rate.

(2) *Repetitive stimuli studies:* To test the hypothesis that successive emotional stimuli may have a cumulative effect on CVP, in the resting patient (i.e., not talking) a series of potentially "stressful" topics for the particular patient being studied were introduced in fairly rapid succession.

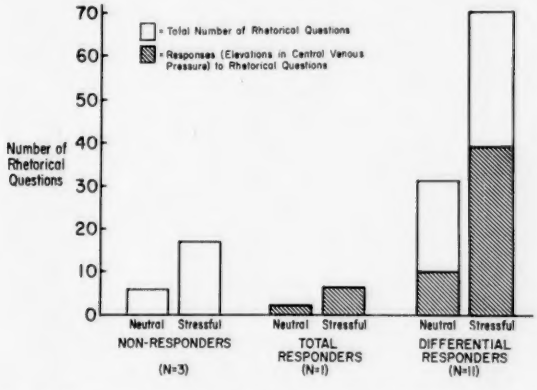


Fig. 5.—Percentage of responses to "stressful" and "neutral" rhetorical questions. Three patients showed no CVP response to any question, one showed a CVP response to all questions, and there were 11 differential responders. The differences in the latter group are significant at the 5 per cent level of confidence.

*Example:* Fig. 6 shows the responses of a 45-year-old farmer with severe rheumatic heart disease, mitral stenosis, tricuspid insufficiency, and auricular fibrillation, Class IV. The patient was an aggressive individual accustomed to much fighting in his younger days. Because of his heart disease, fighting had been restricted in recent years, although he continued to have numerous arguments, particularly with his wife. During the 14 minutes of the recording illustrated, the CVP is observed to increase from 175 to 215 mm. of water and persist at about this level until the forty-eighth minute, when the topic of cardiac surgery was introduced and a further increase to 220 mm. of water occurred. The CVP level subsided rapidly after the patient was encouraged to relax, suggesting that the previous prolonged elevation was related, at least in part, to the interview and probably its content. Respiratory rate, rhythm, and depth were virtually unchanged throughout the recording. Heart rate changed moderately, although the presence of auricular fibrillation makes it difficult to evaluate the relationship between CVP and heart rate. (The effects of arm motion are seen readily in two instances.) Three details of actual tracings (Fig. 7) show the responses to three stimuli regarded as emotionally "stressful."

*Comment:* The cumulative effect of a series of emotionally "stressful" stimuli is seen in these tracings. When it is remembered that the patient is lying quietly, without speaking, the probable dimensions of the load in terms of increased venous return which had been placed on the heart by the original emotional stimuli become apparent.

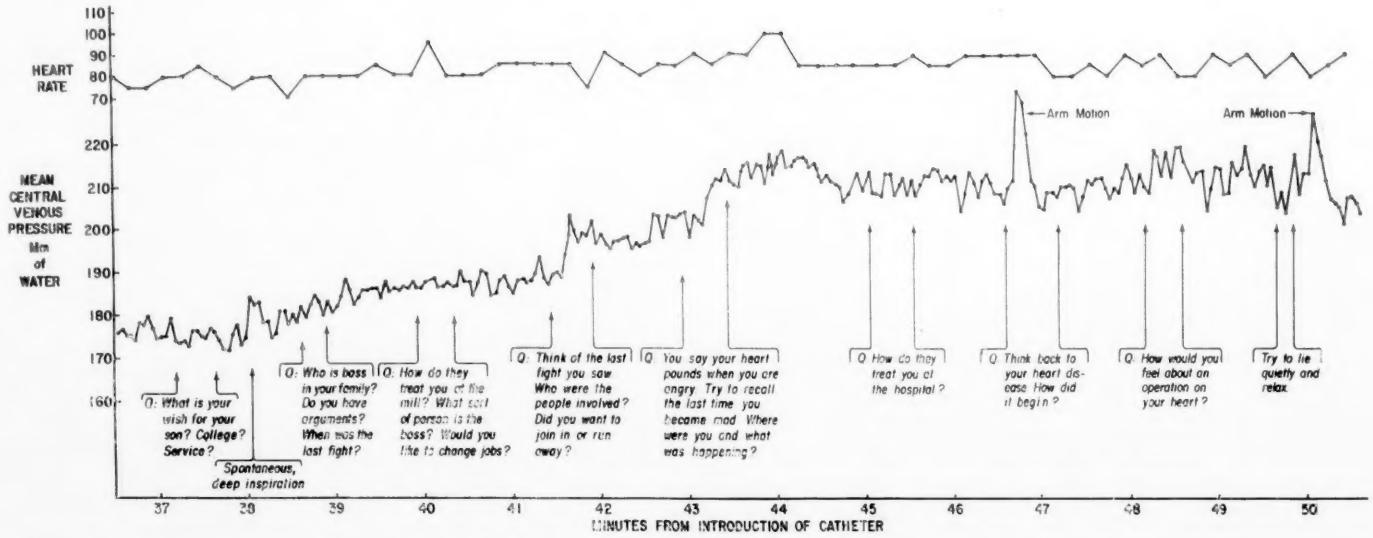
HEART RATE
MEAN CENTRAL VENOUS PRESSURE Mm of WATER
110
100
90
80
70
220
210
200
190
180
170
160
Arm Motion
Arm Motion
Q: What is your wish for your son? College? Service?
Spontaneous, deep inspiration
Q: Who is boss in your family? Do you have arguments? When was the last fight?
Q: How do they treat you at the mill? What sort of person is the boss? Would you like to change jobs?
Q: Think of the last fight you saw Who were the people involved? Did you want to join in or run away?
Q: You say your heart pounds when you are angry. Try to recall the last time you became mad. Where were you and what was happening?
Q: How do they treat you at the hospital?
Q: Think back to your heart disease. How did it begin?
Q: How would you feel about an operation on your heart?
Try to lie quietly and relax.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
MINUTES FROM INTRODUCTION OF CATHETER

Fig. 6.—Central venous pressure (CVP) and heart rate over a 14-minute period in a patient severely ill with rheumatic heart disease, mitral stenosis, tricuspid insufficiency, and auricular fibrillation, Class IV. There is an increase in CVP of 45 mm. of water in response to the cumulative effect of successive "stressful" questions.



VOL
10
195

(3) *Interview studies:* The hypotheses underlying this group of studies were: (a) that elevations in CVP may occur during interviews in association with discussion of disturbing life situations, and (b) that over an extended period of time there may be a cumulative effect on CVP of repeated emotional stimuli in patients with congestive heart failure.

Ten patients were interviewed by the psychiatrist (C.R.V.) after the other physical and rhetorical question studies had been completed. The interviews were tape recorded simultaneously with the physiologic variables, and during these interviews the patients spoke freely, often with considerable expression of feeling.

*Example:* Fig. 8 illustrates a segment of an interview with the psychiatrist. The patient is a 57-year-old white millworker with arteriosclerotic heart disease, Class IV. He was followed regularly for two years and his many problems, as well as his clinical responses, were well known to the investigators. The tracing illustrates a rapid increase in CVP from 140 to 185 mm. of water when a potential fight was described with considerable feeling to the psychiatrist. There is only a minimal change in heart rate.

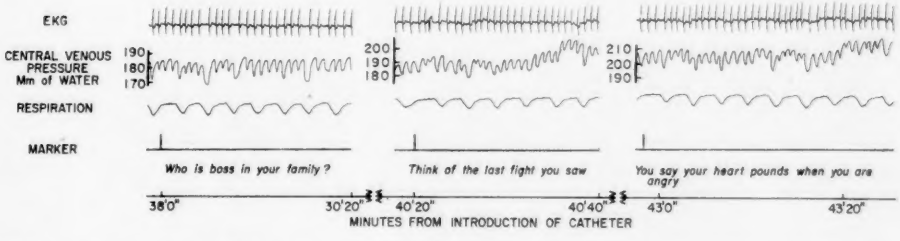


Fig. 7.—Same patient as Fig. 6. Detailed tracings of heart rate, central venous pressure (CVP) and respiration in response to the second, fourth, and fifth rhetorical questions shown in Fig. 6.

*Comment:* Again the increase in CVP in response to emotional stimuli is observed during the interview situation. Such responses may be regarded as simulating to some extent the patient's actual experiences, but are probably considerably less intense since they have been "ventilated" at least once previously during an interview with one of the investigators (D.A.M.) prior to the psychiatric interview in the laboratory.

(4) *Psychiatric follow-up studies:* To obtain further confirmation of the data used in the predictive studies with the "neutral" and "stressful" rhetorical questions, and to evaluate further the roles of emotional and social influences in the lives of patients with congestive heart failure, 30 patients were interviewed by the psychiatrist (C.R.V.). The details of this study are reported elsewhere.[21] Patients were seen initially either on the wards or in the outpatient clinic for one to two hours and, as described above, 10 patients were interviewed in the laboratory. Five patients were followed regularly for from three to sixteen months. In addition to confirming historical data about life situations and experiences, the psychiatrist sought information about the patient's attitudes, feelings, and life conflicts. Special inquiries were made into the ways each patient dealt with both the realistic and symbolic aspects of his interpersonal relationships and with his heart disease.

Emotional reactions may affect the status of a patient's illness in *direct* or *indirect* ways. Reactions which inhibit the patient's effective understanding of his health problems and the regimen prescribed by his physician may be as harmful as no therapy or the wrong therapy. Examples of these *indirect* deleterious emotional influences on the course of the patient's heart disease might include excessive physical exertion, excessive use of salt, or refusal to take medications prescribed by the physician. Fourteen patients (47 per cent) in the group of 30 were found to be benefiting less than might be expected from their prescribed treatment regimens because of such *indirect* factors.

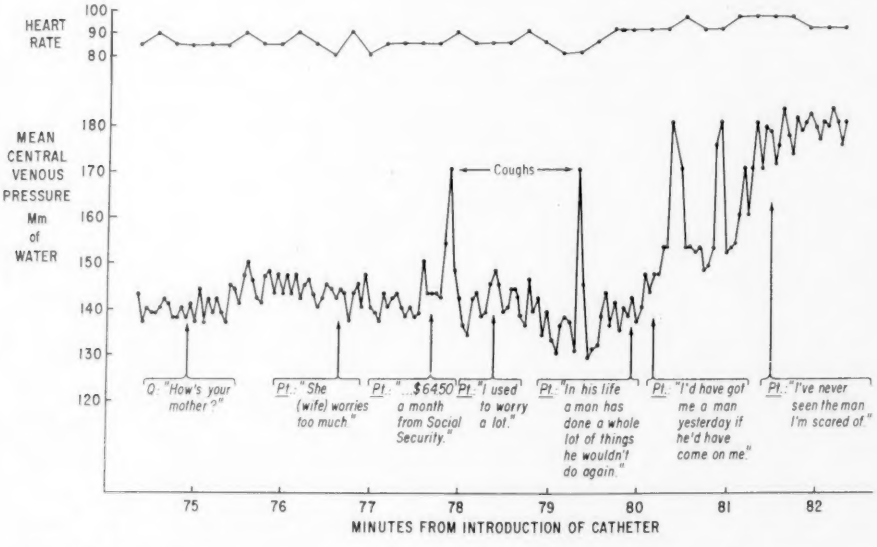


Fig. 8.—Cardiovascular responses of a patient severely ill with arteriosclerotic heart disease, Class IV, during a psychiatric interview. A rapid increase in central venous pressure (CVP) of 45 mm. of water occurs as the patient describes his anger of the previous day. The readily identifiable effects on CVP of two coughs are apparent.

*Example:* A 53-year-old Negro woman with arteriosclerotic heart disease, Class III, developed congestive heart failure abruptly one evening after retiring. This took the form primarily of dyspnea and was diagnosed as pulmonary edema. She remained in cardiac failure for two days before consulting her physician. She responded favorably to digitalis and dietary measures but the improvement was only temporary and she became worse for no apparent reason. Responsibilities at home were pressing, and she developed the idea that if she continued to work conscientiously her symptoms would abate spontaneously. Denying the existence of her illness, she continued to work, with resulting perpetuation of failure and progressive increase of symptoms despite appropriate medical treatment. Examination of her life-long patterns of adaptation revealed that she had always been in excellent physical health and had prided herself on her capacity to work and care for others. Her father died when she was 9 and although heavy work responsibility had been placed on her, she had led a relatively stable life until her husband's death some ten years prior to her present illness. During the same year, both her daughter and mother died, and she was left with complete responsibility for her family's support, a situation to which she outwardly resigned herself. When congestive failure finally developed she was unable to accept her physical disability. Feelings of helplessness and dependency aroused as a result of heart failure were intolerable to her. Not knowing how to modify her activity, she could not

accept her illness and continued in her usual duties even in the face of obvious symptoms of increasing failure. Admission to the hospital, rest, relief of responsibilities at home, and the realistic and firm support of her physician were followed by immediate relief of symptoms without essential change in medication or diet.

*Comment:* Clarification of the patient's life situation, her characteristic behavior patterns, and her areas of emotional conflict made her acceptance of a standard medical regimen possible, and explained in large measure the "intractability" of her congestive failure.

In 19 patients (63 per cent), serious emotional conflicts were regarded as more directly involved in the onset, exacerbation, or perpetuation of congestive failure. These *direct* influences were associated with physiologic disturbances and probably exerted their effects on the cardiovascular system through neural or humoral pathways. In 7 of this group a definite temporal relationship between the disturbing life situations and the onset or exacerbation of failure was established. In another 5 patients probable relationships and in 7 questionable relationships were established. For some of the 19 patients the influences of disturbing experiences on the cardiovascular system were brief, lasting only a few minutes to a few hours. For most, these influences were of longer duration.

Of the 30 patients studied, 24 (80 per cent) showed the effects of chronic emotional strain which seemed to restrict optimum psychologic and physiologic adjustment to the presence of heart disease.

*Example:* A 61-year-old retired high school principal with hypertensive and arteriosclerotic heart disease, Class IV, developed congestive heart failure gradually over a period of months. Medical treatment and rest afforded him complete relief of symptoms. He returned to work, but because of advice to decrease his work load retired as principal and moved to another county to teach in a high school. Although his responsibilities appeared less, the new situation turned out to be more "stressful." Disciplinary and academic problems arising in the intimacy of the classroom were such that he found himself carrying a much greater "load" than he had as principal. The actual work "load" was recognizably less; in contrast, the emotional work "load" for him was extreme. Symptoms of failure returned and he developed increasing dyspnea and peripheral edema. Medical management failed and it was apparent that his cardiac reserve had deteriorated. An interview with him revealed the intense emotional strain resulting from the responsibilities of classroom teaching. In his opinion, this strain more than exceeded the decrease in work "load" resulting from his retirement as school principal. Further, he felt that the move from the county where he had been a principal for many years, had established a reputation, and made many close friends also had been a strain for him. In addition to the apparent direct effect of emotional tension on his cardiovascular status, there was the indirect effect of his persistence in teaching despite his obviously increasing cardiac symptoms. Evaluation of his background confirmed his lifelong inability to accept a role of dependency and his denial of this in the face of progressively serious symptoms.

*Comment:* Knowledge of the patient's personality, his life situation, and the circumstances in which he became ill made it possible to understand the exacerbation of his symptoms in spite of what appeared to be a superior medical regimen.

*Example:* A 48-year-old bachelor with rheumatic heart disease and mitral stenosis, Class III, had made a remarkable adjustment to his progressive heart disease. Repeated attacks of rheumatic fever began at the age of 10 years. He had been able to accept the need for frequent modi-

fication in his life adjustment as his cardiac reserve decreased. Reared as the son of a farmer, he started work in this occupation, but as his heart disease progressed he took less demanding jobs. Eventually he reached a point where the simple exertion required of a filling station attendant became excessive for him. He was hospitalized to evaluate the feasibility of mitral commissurotomy.

After review of his experiences, it was discovered that his heart remained relatively well compensated by avoidance of physical exertion and provided unexpected or unusual emotional experiences did not occur. For example, on one occasion he was in a serious automobile accident, but was unhurt. No dyspnea or discomfort occurred at the time of the accident, but shortly afterward he noted increasing dyspnea which persisted during the night and correlated with his progressive anxiety as he ruminated about the accident.

*Comment:* An intense, but acute, emotional experience was associated with the development of symptoms in this patient, but only after he had perceived and reacted to the event and its full implications. It is the patient's psychophysiologic response to a symbolically significant experience which is important, and not necessarily the isolated event itself.

VOL. 10 1959

## PERIPHERAL VENOUS PRESSURE OBSERVATIONS

Venous pressure reflects the net contribution of a number of cardiovascular and extracardiac influences and the mechanisms by which it is altered may vary with different types of heart disease, with varying degrees of cardiac compensation, and in response to differing functional demands on the cardiovascular system. One possible contribution to increases in CVP might stem from increases in peripheral venous tone in response to emotional stimuli. Accordingly, under controlled conditions, a study was made of the effects of specific emotional stimuli on the perpheral venous tone of volunteers with ostensibly normal cardiovascular systems.

*Subjects.*—Six paid university students, aged 20 to 25 years, with prominent right forearm veins free of tributaries were chosen from healthy volunteers. Following the attachment of the recording instruments and the instruction of the patient as described previously, the vein occluders were applied and a base line recorded (Fig. 2). Allowing sufficient time between tests to permit stabilization of the base line, each of the following seven procedures was carried out for 30 seconds: (1) rapid shallow breathing (panting); (2) simple mental arithmetic, i.e., "3 multiplied by 65," considered to be a uniformly "neutral" stimulus for these students; (3) difficult mental arithmetic, i.e., "73 multiplied by 49," considered to be a uniformly "stressful" stimulus for this particular group and one which was intensified for all of the subjects by offering additional payment if the correct answer was obtained in 30 seconds; (4) constant tensing of the right forearm and hand muscles; (5) slow deep breathing; (6) finger motion of the right hand; (7) observation of card No. 13 MF of the Thematic Apperception Test (TAT), which was considered to be a variable stimulus.[22] This card is an ambiguous drawing open to different individual interpretations believed to be related to each individual's unconscious emotional conflicts. The card should theoretically evoke individual emotional responses of varying degrees and durations.

*Results.*—The data from this study are reported in detail elsewhere.[23] The difference between the mean increase for difficult arithmetic and the increases associated with any of the other five test procedures (excluding the THT card test) was highly significant ($P < 0.001$) (Fig. 9). The mean increase in response to this procedure for the 6 subjects was 26 per cent above the average base-line level of 324 mm. of water. When the mean increases were determined planimetrically as an integrated function of both the height and duration of each pressure curve, the response to difficult mental arithmetic was again much greater than those to the other five stimuli ($P < 0.001$). There were no significant differences between the means of the remaining five procedures when the two indices of height and duration of response were employed.

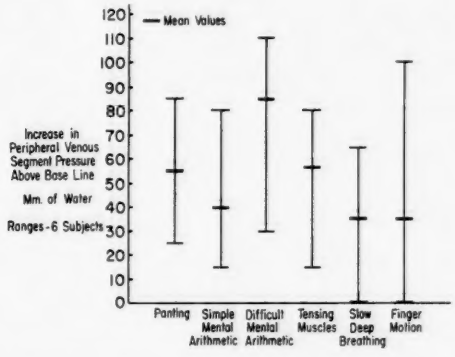


Fig. 9.—Mean and ranges of increases in peripheral venous segment pressure (PVSP) in 6 subjects in response to physical and emotional stimuli.

*Example:* Fig. 10 shows the responses of one volunteer to both simple arithmetic and difficult mental arithmetic. The persistence of the increased PVSP after the subject has been instructed to discontinue his mental efforts suggests the intensity of both the stimulus and the increased venous tone. There is little or no alteration in respiration, and, although there is an increase in heart rate of from 5 to 20 per minute, this can have little direct influence on events within the isolated vein segment.

*Example:* Fig. 11 illustrates an isolated peripheral vein segment of a patient with arteriosclerotic heart disease, Class II. The patient was untreated at the time of the study. Failure developed following his arrest for the murder of his son whom he had shot accidentally after the latter attacked him during a drinking spree. When asked to think about the outcome of the trial there was a marked drop in the psychogalvanic skin response and a gradual increase in PVSP of about 30 mm. of water in the next two minutes. Neither heart rate nor periodic breathing, which had been present throughout most of the tracing, changed materially.

*Comment:* The substantial effects of emotional stimuli on peripheral venous segment tone as measured by pressure changes in 6 normal volunteers were of greater magnitude and duration than those seen with physical stimuli. The one patient with congestive heart failure whose PVSP has been studied so far showed similar responses.

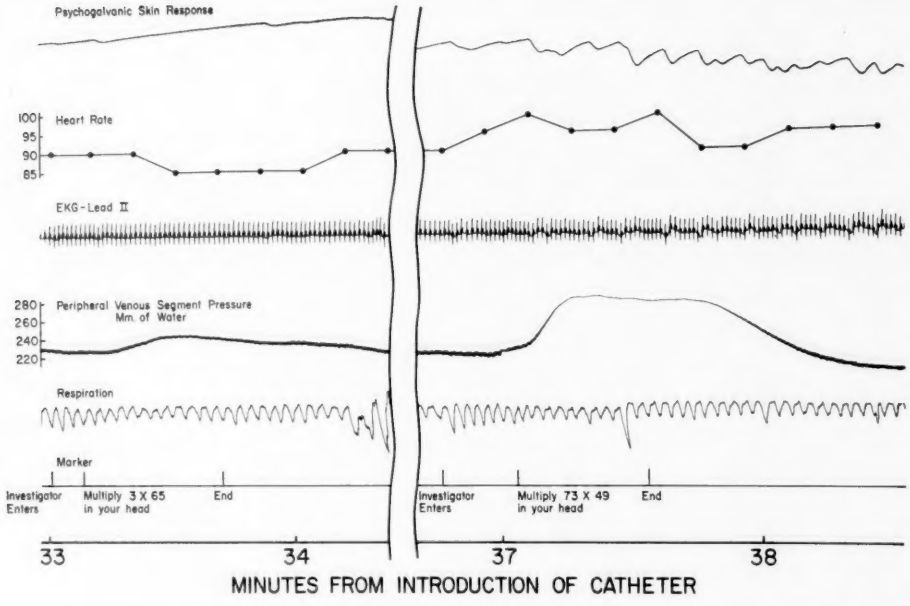


Fig. 10.—Comparison of the peripheral venous segment pressure (PVSP), heart rate, respiration, and psychogalvanic skin response in a normal volunteer to simple and difficult mental arithmetic.

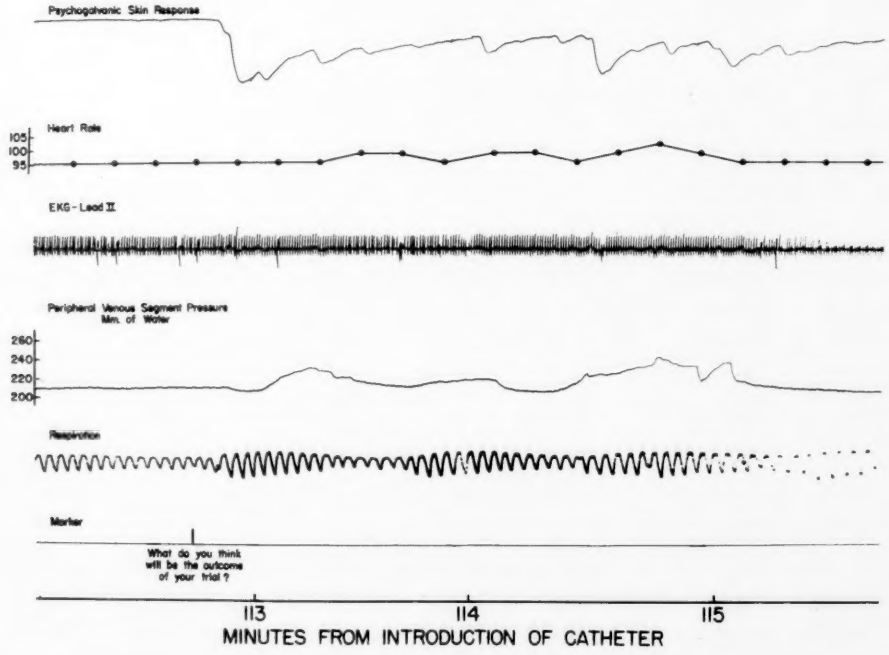


Fig. 11.—Psychogalvanic skin response, heart rate, ECG-Lead II, peripheral venous segment pressure (PVSP), and respiration in a patient with arteriosclerotic heart disease, Class II, who was asked a rhetorical question about the outcome of his trial for murder, an event which appeared to be temporally associated with the onset of his symptoms of congestive heart failure.

DISCUSSION

The data presented tend to support the hypothesis advanced by Chambers and Reiser[3] that there is an association between a patient's life situations, his emotional reactions to these, and the onset, exacerbation, or perpetuation of congestive heart failure. These influences may be indirect deleterious modifications of behavior or attitudes, or they may be direct influences operating through physiologic mechanisms. These kinds of data might offer at least partial explanations beyond those afforded by the progressive pathologic changes associated with the underlying organic diseases for the insidious onset of congestive failure in half the patients studied by Sodeman and Burch[1] and Boyer, Leach, and White.[2] The potential contribution to the therapeutic regimen of data about the social and emotional circumstances under which patients develop congestive failure is discussed in another communication.[21]

The data presented here tend to support an association, at least in some patients, between emotional reactions in response to specific life situations and increases in central venous pressure. The increases in central venous pressure may be the net expression of several physiologic mechanisms, including changes in heart rate, arterial blood pressure, respiratory rate, intrathoracic pressure, rapid increase in blood volume, increased muscle tension, or alterations in tone of central and/or peripheral veins. These central venous pressure elevations tend to be both higher and of longer duration in the patients with more serious decompensation, and in association with those emotional reactions which patients appear to experience most intensely.

Burch,[24,25] and Wood, Litter, and Wilkins[26] demonstrated that the tone of peripheral veins is increased in patients with congestive heart failure and that it decreases with improvement in cardiac compensation and with the administration of hexamethonium. Halmagyi, Felkai, Ivanyi, and Hetenyi[27] studied the effects of ergotamine derivatives, sodium nitrite, tetra-ethylammonium bromide and dibenamine on central venous pressure and concluded that the sympathetic nervous system plays a substantial role in maintaining venous hypertension during cardiac failure. Ross, Wilson, and Hickam[28] have reported that physically normal psychiatric patients whose motor responses had been "virtually eliminated" by succinyl choline showed substantial increases in CVP in response to strong stimulation of the autonomic nervous system by electric shock therapy. Brief stimulation produced mean CVP increases of 69 mm. of water, and longer stimulation produced mean elevations of 169 mm. of water, in some cases sustained for as long as 5 minutes. These data support the hypothesis that autonomic stimulation may have profound effects on CVP even in subjects with normal cardiovascular systems.

At least one possible factor contributing to alterations in CVP may be PVSP changes in response to emotional as well as physical stimuli. A detailed presentation of these data will be found elsewhere.[23] It may be postulated that peripheral venous constriction is at least partially mediated through a reflex because of the rapidity with which the phenomenon occurs in a virtually isolated vein segment. In addition, the possibility remains that humoral substances or small quantities

of arterial blood reaching the segment through the vasa vasorum may also contribute.

Landis and Hortenstine[29] estimated that 66 to 75 per cent of the blood volume is contained within the venous system. The functional implications of this distribution of available blood volume in relation to venous return and the regulation of cardiac output have been reviewed at length recently.[30,31,32] The quantitative implications of vasomotor activity on the venous side of the circulation have been considered by Eckstein and Hamilton.[33] These authors calculated that small infusions of epinephrine and levarterenol produced peripheral venous constriction which promptly returned an estimated 500 c.c. of blood to the central veins from dependent extremities. This blood is available for the maintenance of cardiac output in the healthy individual, but the addition of as much as 500 c.c. to an estimated central blood volume of 1,200 c.c.[34] could represent a substantial load for a poorly compensated heart. If the latter is unable to propel this increased volume of blood, and if the central venous reservoirs are maximally distended,[29,35] both central and peripheral mean venous pressure may be elevated significantly.

In addition to the effects of venous tone, and of central venous pressure on peripheral venous pressure, the latter is influenced by total blood volume, the rate of blood flow through the veins, by compression of the veins by surrounding tissues,[36] and by the hydrostatic pressure exerted by the column of blood between the peripheral and central veins. If, as a result of the increased amount of blood shifted rapidly toward the diseased heart by emotional influences, acting reflexly or humorally on the peripheral veins, peripheral venous pressure increases, edema may develop.

Landis, Brown, Fauteux, and Wise[37] showed that, in animals with decreased cardiac "competence," venous pressure was normal when the animals were at rest, but increased above the limits of normal when the animals were exercised. Landis and Gibbon[38] showed that peripheral venous pressure in the human forearm, at levels of water of 100 mm. or greater, was accompanied by filtration of fluid into the extravascular spaces at the rate of 0.0033 ml. per minute per 100 ml. of forearm tissue per centimeter of water elevation in venous pressure. According to these estimates an increase in peripheral venous pressure of 100 mm. of water above normal could result in the filtration of 250 to 300 ml. of edema fluid in a ten-minute period. Reichsman and Grant[39] observed in 3 patients that as cardiac decompensation developed, following withdrawal of digitalis, the rise in peripheral venous pressure preceded weight gain and edema formation. These data are somewhat at variance with Warren and Stead's[40] report of 2 patients with recently compensated hearts in whom weight gain preceded the rise in venous pressure when excess sodium chloride was given.

Most of the available data are consistent with the view that emotional, as well as physical factors, intermittently or recurrently, may return more blood to the failing heart than it can propel, with a consequent gradual distention of the central veins and eventually an increase in pressure in all veins, but particularly the more permeable veins of the liver, leading to the accumulation of ascites and edema.

The Landis hypothesis holds that frequent increases in central venous pressure in patients with congestive heart failure, from whatever cause, could result in accumulation of fluid in the extravascular space. The quantity of fluid which accumulates will depend on the height and duration of central venous pressure and the previous degree of failure. Experimental observations by Ball, Davis, and Goodkind[41] also suggest that one of the initial events leading to congestive heart failure may be an elevated venous pressure. These investigators produced ascites and edema experimentally both in dogs with thoracic vena cava constriction, and in dogs with pulmonary artery constriction. Although sodium intake was restricted, venous pressure rose and there was extravasation of fluid, increased excretion of aldosterone, and sodium and water retention.

Much further work is required to clarify the precise nature and sequence of physiologic events which lead to congestive heart failure, but it remains within the realm of possibility that initial events on the venous side may be of primary importance. Whether of primary importance or only as contributing factors, the data presented here support the hypothesis that disturbing life situations and the emotional reactions they evoke may produce recurrent as well as cumulative increases in peripheral and central venous pressure. These, in turn, may result in loculation of blood in the readily distensible veins of the thorax and abdomen with the eventual production of edema and the symptoms of congestive heart failure.

VOL.
10
1959

## SUMMARY

1. The literature dealing with the epidemiologic aspects of congestive heart failure and with the influences of emotional factors on certain aspects of cardiovascular function, including congestive heart failure (CHF), is reviewed briefly.

2. Examples of central venous pressure (CVP) increases in response to physical stimuli, to "neutral" and "stressful" rhetorical questions, and to repeated emotional stimuli, in 25 studies in 22 patients, and to psychiatric interviews in 10 patients are reported. Emotionally charged questions which can evoke increases in CVP can be predicted from brief "medical" interviews at the 5 per cent level of confidence. Increases in CVP of as much as 45 mm. of water in response to the cumulative effect of successive emotionally charged questions are reported.

3. Psychiatric studies of 30 patients with CHF documenting the direct (63 per cent) and indirect (47 per cent) deleterious influences of emotional reactions on CHF are discussed.

4. Peripheral venous segment pressure (PVSP) changes in response to certain physical and emotional stimuli in 6 healthy subjects are reported. "Difficult" mental arithmetic produced highly significant increases in PVSP when compared to the physical stimuli and to "simple" mental arithmetic.

5. The possible effects of emotional influences on peripheral venous tone as it may modify venous return and the regulation of cardiac output are discussed. The data tend to support Landis' concept of the mechanism by which congestive heart failure may develop insidiously in some patients.

REFERENCES


1. Sodeman, W. H., and Burch, G. E.: The Precipitating Causes of Congestive Heart Failure, Am. Heart J. **15**:22, 1938.
2. Boyer, N. H., Leach, C. E., and White, P. D.: Underlying Causes and Precipitating Factors of Congestive Heart Failure, in Blood, Heart and Circulation, Publication 13, American Association for the Advancement of Science, Lancaster, 1940, The Science Press, pp. 203-212.
3. Chambers, W. N., and Reiser, M. F.: Emotional Stress in the Precipitation of Congestive Heart Failure, Psychosom. Med. **15**:38, 1953.
4. Stevenson, I. P., Duncan, C. H., Wolf, S., Ripley, H. S., and Wolff, H. G.: Life Situations, Emotions and Extrasystoles, Psychosom. Med. **11**:257, 1949.
5. Wolf, S. G., Cardon, P. V., Jr., Shepard, E. M., and Wolff, H. G.: Life Stress and Essential Hypertension, Baltimore, 1955, Williams & Wilkins Company, pp. 17-30.
6. Ring, A., and Blankfein, J.: Paroxysmal Ventricular Tachycardia in an Apparently Normal Heart, Ann. Int. Med. **42**:680, 1955.
7. Harvey, W. P., and Levine, S. A.: Paroxysmal Ventricular Tachycardia Due to Emotion; Possible Mechanism of Death From Fright, J.A.M.A. **150**:479, 1952.
8. Grollman, J.: Physiological Variations in the Cardiac Output of Man, Am. J. Physiol. **89**:584, 1929.
9. Hickam, J. B., Cargill, W. H., and Golden, A.: Cardiovascular Reactions to Emotional Stimuli, Effect on the Cardiac Output, Arteriovenous Oxygen Difference, Arterial Pressure and Peripheral Resistance, J. Clin. Invest. **27**:290, 1948.
10. Stead, E. A., Jr., Warren, J. V., Merrill, A. J., and Brannon, E. S.: The Cardiac Output in Male Subjects as Measured by the Technique of Right Atrial Catheterization: Normal Values With Observations on the Effects of Anxiety and Tilting, J. Clin. Invest. **24**:326, 1945.
11. Stevenson, I. P., Duncan, C. H., and Wolff, H. G.: Circulatory Dynamics Before and After Exercise in Subjects With and Without Structural Heart Disease During Anxiety and Relaxation, J. Clin. Invest. **28**:1534, 1949.
12. Duggan, J. J., Love, V. L., and Lyons, R. H.: A Study of Reflex Venomotor Reactions in Man, Circulation **7**:869, 1953.
13. Page, E. G., Hickam, J. B., Sieker, H. O., McIntosh, H. D., and Pryor, W. W.: Reflex Venomotor Activity in Normal Persons and in Patients With Postural Hypotension, Circulation **11**:262, 1955.
14. Burch, G. E., and Murtadha, M.: A Study of the Venomotor Tone in a Short Intact Venous Segment of the Forearm of Man, Am. Heart J. **51**:807, 1956.
15. Salzman, E. W.: Reflex Peripheral Venoconstriction Induced by Carotid Occlusion, Circulation Res. **5**:149, 1957.
16. Alexander, R. S.: Participation of the Venomotor System in Pressor Reflexes, Circulation Res. **2**:405, 1954.
17. Burch, G. E., and Winsor, T.: Phlebomanometer; a New Apparatus for Direct Measurement of Venous Pressure in Large and Small Veins, J.A.M.A. **123**:91, 1943.
18. Doupe, J., Krynauw, R. A., and Snodgrass, S. R.: Some Factors Influencing Venous Pressure in Man, J. Physiol. **92**:383, 1938.
19. Darrow, C. W.: The Galvanic Skin-Reflex and Finger Volume Changes, Am. J. Physiol. **88**:219, 1929.
20. Sabshin, M., Hamburg, D. A., Grinker, R. R., Persky, H., Basowitz, H., Korchin, S. J., and Chevalier, J. A.: Significance of Preexperimental Studies in the Psychosomatic Laboratory, A. M. A. Arch. Neurol. & Psychiat. **78**:207, 1957.
21. Vernon, C. R., Martin, D. A., and White, K. L.: Management of Patients With Congestive Heart Failure: A Psychophysiologic Approach, J.A.M.A. In press.
22. Murray, H. A.: Thematic Apperception Test Manual, Cambridge, Mass., 1943, Harvard University Press, 20 pp.
23. Martin, D. A., White, K. L., and Vernon, C. R.: Influence of Emotional and Physical Stimuli on Pressure in the Isolated Vein Segment, Circulation Res. In press.
24. Burch, G. E.: A Method for Measuring Venous Tone in Digital Veins of Intact Man. Evidence for Increased Digital Venous Tone in Congestive Heart Failure, Arch. Int. Med. **94**:724, 1954.
25. Burch, G. E.: Evidence for Increased Venous Tone in Chronic Congestive Heart Failure, Arch. Int. Med. **98**:750, 1956.
26. Wood, J. E., Litter, J., and Wilkins, R. W.: Peripheral Venoconstriction in Human Congestive Heart Failure, Circulation **13**:524, 1956.
27. Halmagyi, D., Felkai, B., Ivanyi, J., and Hetenyi, G., Jr.: The Role of the Nervous System in the Maintenance of Venous Hypertension in Heart Failure, Brit. Heart J. **14**:101, 1952.
28. Ross, J. C., Wilson, W. P., and Hickam, J. B.: Reflex Venoconstrictor Response to Strong Autonomic Stimulation, Clin. Res. **6**:399, 1958.


VOL 10 195

29. Landis, E. M., and Hortenstine, J. C.: Functional Significance of Venous Blood Pressure, Physiol. Rev. **30**:1, 1950.
30. Brecher, G. A.: Venous Return, New York, 1956, Grune & Stratton, Inc.
31. Freis, E. D., and Rose, J. C.: Sympathetic Venous System, Vascular Volume and Venous Return in Relation to Cardiovascular Integration, Am. J. Med. **22**:175, 1957.
32. Guyton, A. P., Lindsey, A. W., Abernathy, B., and Langston, J. B.: The Mechanism of the Increased Venous Return and Cardiac Output Caused by Epinephrine, Am. J. Physiol. **192**:126, 1958.
33. Eckstein, J. W., and Hamilton, W. K.: The Pressure-Volume Responses of Human Forearm Veins During Epinephrine and Nor-epinephrine Infusions, J. Clin. Invest. **36**:1663, 1957.
34. Doyle, J. T., Wilson, J. S., Lepine, C., and Warren, J. V.: An Evaluation of the Measurement of Cardiac Output and the So-Called Pulmonary Blood Volume by the Dye-Dilution Method, J. Lab. & Clin. Med. **41**:29, 1953.
35. Alexander, R. S., Edwards, W. S., and Ankeney, J. L.: Distensibility Characteristics of the Portal Vascular Bed, Circulation Res. **1**:271, 1953.
36. Ryder, H. W., Molle, W. E., and Ferris, E. B., Jr.: Influence of the Collapsibility of Veins on Venous Pressure, Including a New Procedure for Measuring Tissue Pressure, J. Clin. Invest. **23**:333, 1944.
37. Landis, E. M., Brown, E., Fauteux, M., and Wise, C.: Central Venous Pressure in Relation to Cardiac "Competence," Blood Volume and Exercise, J. Clin. Invest. **25**:1237, 1946.
38. Landis, E. M., and Gibbon, J. H., Jr.: The Effects of Temperature and Tissue Pressure on the Movement of Fluid Through the Human Capillary Wall, J. Clin. Invest. **12**:105, 1933.
39. Reichsman, F., and Grant, H.: Some Observations on the Pathogenesis of Edema in Cardiac Failure, Am. Heart J. **32**:438, 1946.
40. Warren, J. V., and Stead, E. A., Jr.: Fluid Dynamics in Chronic Congestive Heart Failure, Arch. Int. Med. **73**:138, 1944.
41. Ball, W. C., Jr., Davis, J. O., and Goodkind, M. J.: Ascites Formation Without Sodium Intake in Dogs With Thoracic Inferior Vena Cava Constriction and in Dogs With Right-Sided Congestive Heart Failure, Am. J. Physiol. **188**:578, 1957.

VOL.
10
959

# SOME METHODOLOGIC PROBLEMS IN THE LONG-TERM STUDY OF CARDIOVASCULAR DISEASE: OBSERVATIONS ON THE FRAMINGHAM STUDY

TAVIA GORDON,*
FELIX E. MOORE,**
DEWEY SHURTLEFF,*** AND
THOMAS R. DAWBER, M.D.,***
BETHESDA, MD.

(Received for publication Feb. 18, 1959)

VOL.
10
195

## INTRODUCTION

IN 1948, the United States Public Health Service undertook to study the factors associated with development of atherosclerotic and hypertensive cardiovascular disease by long-term surveillance of a sample of the adult population of Framingham, Massachusetts. The study has been reported at various stages,[1-3] but it may be useful to reconsider at this time the reasons for bringing the study into being and to examine how far it has proceeded toward its stated objectives. We will consider only a few of the many problems that have arisen, but try to put them in the context of more general considerations.

It was recognized in 1948, and it is even clearer in the light of all of the research of the succeeding 10 years, that atherosclerotic and hypertensive heart diseases typically develop slowly and may be present for many years before unequivocal clinical signs can be observed. In 1948 most investigators studying these diseases agreed that the precise etiologies were unknown, and this remains essentially the situation 10 years later, although the research of the past decade has contributed to the development of several new and highly plausible hypotheses. It was also felt that the onset of these diseases was probably influenced by many factors, and that these would prove difficult to disentangle.

In the light of these considerations, it was decided to inaugurate a study in an American community from which a sample of the adult population would be drawn and characterized with respect to clinically detectable atherosclerotic or

---

All of the authors are or have been members of the National Heart Institute, Public Health Service, which has administrative responsibility for the Framingham Study.

*National Cancer Institute.

**University of Michigan.

***National Heart Institute.

hypertensive cardiovascular disease. The cohorts of persons free of clinically detectable disease would then be followed by means of clinical examinations at 2-year intervals. With the progression of disease in the cohorts, it was believed that comparison of those developing clinical signs of disease with those remaining free of disease during a lengthy follow-up period would disclose associations and permit inferences about etiology. Originally it was intended to follow only the "normal" cohorts,[1] but subsequently it was decided to include all of the originally examined group in the routine of re-examination.

It was emphasized at the outset of the study that the town of Framingham could not in any statistical sense be considered a sample of a definable universe. Reasons for the selection of this community have been previously described[1] and are therefore reviewed here only briefly. (a) Because of the proposal to examine the study cohorts in a clinic and to carry out follow-up by means of clinic examinations, it was felt necessary to restrict the study to a community with a sizable adult population, but compact enough to allow easy access to the central clinic. (b) The community should be relatively self-sufficient with respect to all medical services (including hospitalization), so that there could be ready flow of information between the medical community and the clinic. (c) At the same time the clinic should be located reasonably close to a center of medical education and research from which a consultant staff could be drawn and where the resident staff could receive additional training. (d) There should be evidence that the resources of community organization could be drawn upon to ensure maximum participation. (e) And finally, after due regard to all these factors, the community should not be "atypical"—a most difficult qualification to describe in any objective terms.

Framingham satisfied all of these conditions. Furthermore, it had been the site of a pioneer community health demonstration in the field of tuberculosis during the years 1917 to 1923, which had been well received in the community. Investigations by the staff director and the health educator assigned to the project confirmed the impression of continued community cooperation.

## THE STUDY SCHEME

Based on the 1949 and 1950 lists of residents for the town, a sample of the adult Framingham population aged 30-59 was selected for study. A systematic sampling scheme was used after "family" groupings (consisting of all individuals with the same surname living at the same address and within the eligible age range on Jan. 1, 1950) were arranged, first, by precinct of residence (eight precincts), then by size of "family" grouping, and finally in alphabetical order. Two out of every three "family" groupings were selected.

A clinic was set up near the Framingham Union Hospital and the sample group was invited to the clinic for examination. By this means study cohorts could be characterized with respect to age, sex, other demographic characteristics, and a large variety of clinical and physiologic characteristics. Every 2 years they would be re-examined at the clinic. During this follow-up some originally free of coronary or hypertensive heart disease would develop the disease. Data would then be available to test a set of specific hypotheses about conditions

predisposing to the development of atherosclerotic or hypertensive cardiovascular disease; e.g., that elevations in serum cholesterol, blood pressure, or relative weight were associated with a higher incidence of coronary heart disease. Furthermore, the development of cardiovascular disease could be placed in the context of changes occurring in the total study group, such as shifts in blood pressure, serum cholesterol levels, or weight.

SAMPLE RESPONSE

It was felt that a study group of 6,000 was about the largest that could be investigated with the resources available. Only rough estimates of the incidence of cardiovascular disease were available, but these suggested that a group of this size in the age range 30-59 would provide a sufficient yield of new disease for analytical purposes. Considerations on which estimates of sample size were based are described in an earlier report.[1]

Volunteer workers were used and all available resources of the community were employed to stimulate interest and cooperation, including at the outset a blanket invitation to any town resident in the age range 20 to 70 to come to the clinic for examination. Initial expectations, based on the great rush of volunteers at the beginning, was for a nonresponse not in excess of 10 per cent, so a sample roster of approximately 6,600 men and women who were in the age range 30 to 59 on Jan. 1, 1950, was drawn. As originally listed the sample consisted of 6,587 names, which was reduced to 6,510 after correction for duplication of names and removal of persons found to be outside the eligible age limits. The number 6,510 appears in a report published in 1957.[3] Since then, three more duplicates have been found in the nonrespondent group, thus reducing the total drawn sample to 6,507. In the end, only 4,494 or 68.8 per cent of the drawn sample came to the study clinic for examination in spite of many and varied appeals to nonrespondents. (Table I.)

For women in each of the age groups under 45 the response rate was around 75 per cent; for men in these age groups it was around 66 per cent. For men or women aged 55 or over the response rate was around 60 per cent. Furthermore, response varied according to the section of town which presumably means that it varied according to socioeconomic characteristics. From one precinct to another response ranged from 57 to 73 per cent for men and from 64 to 76 per cent for women.

The volunteer community workers who made all of the contacts with the persons drawn in the sample were asked to secure reasons for nonresponse. The reasons they reported are of uncertain validity.[2] However, before the end of the initial examination period, 6.5 per cent of the persons in the sample had moved from Framingham and 1.1 per cent were dead (Table II). Of the 6,015 persons in the sample alive and resident in Framingham at the end of the initial examination period, 74.3 per cent had been examined in the clinic. The initial examinations, which were originally planned for a much shorter period, extended from September, 1948, through August, 1952, with a handful of respondents coming in even later.

TABLE I. RESPONSE BY AGE AND SEX
(Framingham Study: Drawn Sample)

| AGE AND SEX | TOTAL | EXAMINED | NOT EXAMINED | PER CENT EXAMINED |
|---|---|---|---|---|
| Both sexes | 6,532 | 4,494 | 2,038 | 68.8 |
| Men | 3,086 | 2,036 | 1,050 | 66.0 |
| 30-34 | 628 | 417 | 211 | 66.4 |
| 35-39 | 529 | 385 | 144 | 72.8 |
| 40-44 | 540 | 358 | 182 | 66.3 |
| 45-49 | 461 | 301 | 160 | 65.3 |
| 50-54 | 481 | 314 | 167 | 65.4 |
| 55-59 | 447 | 261 | 186 | 58.4 |
| Women | 3,446 | 2,458 | 988 | 71.3 |
| 30-34 | 670 | 506 | 164 | 75.5 |
| 35-39 | 625 | 466 | 159 | 74.6 |
| 40-44 | 576 | 438 | 138 | 76.0 |
| 45-49 | 566 | 391 | 175 | 69.1 |
| 50-54 | 538 | 367 | 171 | 68.1 |
| 55-59 | 471 | 290 | 181 | 61.6 |

Age is the "declared age" in the lists from which the sample was drawn. Included in these totals are 25 respondents who were found to be out of the age range on examination and who for that reason were dropped from the study group.

TABLE II. RESPONSE AND NONRESPONSE BY REASON
(Framingham Study: Drawn Sample)

| RESPONSE AND REASON | NUMBER OF PERSONS | PERCENT OF | |
|---|---|---|---|
| | | DRAWN SAMPLE | NONRESPONDENTS |
| Total | 6,532 | 100.0 | ——— |
| Examined | 4,494 | 68.8 | ——— |
| Not examined | 2,038 | 31.2 | 100.0 |
| Moved from Framingham | 426 | 6.5 | 20.9 |
| Died | 74 | 1.1 | 3.6 |
| Ill or incapacitated | 74 | 1.1 | 3.6 |
| Refused | 1,464 | 22.4 | 71.8 |

See footnote to Table I. These data differ slightly from those previously published.

It seems clear that there was differential response to the initial examination according to health status. Since participation required a visit to the clinic, the moribund and bedridden were included among the nonrespondents. The reported 1 per cent of the sample (74 persons) who were nonrespondent on this account seems a reasonable estimate for this age group. For the age group 45 to 64 the U.S. National Health Survey reported a figure of 1 per cent for "confined to house" and 1 per cent for "unable to get around alone" in the second quarter of 1957.[4]

Because of the structure of community organization about which the Framingham study was originally built, which included complete dependence on volunteer canvassers, there was no systematic attempt to secure reports on on health status. Except for a small number of refusals attributed to incapacity by the canvassers and a scattering of information from relatives in the study, there is no direct evidence available on the health status of nonrespondents.

If we were to repeat this study we would take all possible steps to secure some minimal data on reported health status and past utilization of medical services by means of interviews in the home by trained interviewers *in advance* of invitation to examination. With this information some estimate of the relation of assumed health status to the refusal rate could be secured. This is the procedure which has been followed in several subsequent studies.[5-7]

In the absence of more direct evidence on the health status of nonrespondents we are left dependent on the evidence of subsequent mortality for inferences concerning differentials in the health status of respondents and nonrespondents to the initial examination. Not all of the mortality information is usable, however. Of the 1 per cent of the sample not examined at the clinic who died during the period between September, 1948, and December, 1952 (54 men and 23 women), it is not now possible to establish with certainty how many were alive at the date when they should normally have been examined and how many had died before that date. Consequently, it does not appear that there is any appropriate basis for comparing this mortality with the 34 deaths among respondents (25 men and 9 women) which occurred during the same period.

If we were to set up another study of this sort in which the period over which the initial cohort was to be examined was a lengthy one, we would insist that an examination date be set up in advance for each individual in the cohort and that on or very near that date the status of the prospective examinee be recorded in at least the following terms: Examination completed, examination refused, person moved prior to examination date, and person dead prior to examination date.

Lacking this kind of information, the earliest mortality experience cannot be used for a comparison of respondent with nonrespondent. However, it seems reasonable to compute mortality rates for the years 1951 and 1952 alone, since the disinclination or incapacity to respond becomes less ambiguous toward the end of the period of convassing. For the nonexamined group the mortality during this period was double that for the examined group (Table III). For men, this contrast was equally evident for both cardiovascular and for noncardiovascular deaths. For women this is not true; none of the deaths in the nonexamined group were from cardiovascular causes.

The strength of the health differential between respondents and the various classes of nonrespondents may be judged from subsequent mortality experience (Table IV). As expected, the level of mortality varies according to the reason given for nonresponse. It was least for those who had moved and greatest for those incapacitated. What is of more moment, there is as yet no clear indication that mortality in the respondent and nonrespondent groups is converging. This is rather surprising. If those who failed to join the study because of illness had

TABLE III. DEATHS FROM SPECIFIED CAUSES ACCORDING TO RESPONSE AND SEX
(Framingham Study: Drawn Sample, 1951-1952)

| SEX AND CAUSE* | TOTAL | EXAMINED | NOT EXAMINED |
|---|---|---|---|
| Men | 52 | 22 | 30 |
| Cardiovascular diseases | 29 | 13 | 16 |
| Cancer | 11 | 3 | 8 |
| Accidents | 3 | 1 | 2 |
| All other | 9 | 5 | 4 |
| Women | 17 | 7 | 10 |
| Cardiovascular diseases | 4 | 4 | – |
| Cancer | 9 | 3 | 6 |
| Accidents | 1 | — | 1 |
| All other | 3 | — | 3 |

*The cause groups are defined by category numbers of the *Sixth Revision of the International List*. Cardiovascular diseases are List numbers 330-334, 400-468; Cancer, List numbers 140-205; Accidents, List numbers E800-E962.

For men the annual death rates per 1,000 population for 1951-1952 were 6.6 (examined) and 11.1 (not examined); for women they were 1.7 (examined) and 3.7 (not examined).

TABLE IV. DEATHS AND AGE-ADJUSTED DEATH RATES BY SEX: SPECIFIED GROUPS

| AGE SPAN, TIME PERIOD, AND POPULATION GROUP | NUMBER OF DEATHS | | AGE-ADJUSTED DEATH RATE | |
|---|---|---|---|---|
| | MEN | WOMEN | MEN | WOMEN |
| Ages 30-59 years, 1950-1952 | | | | |
| Massachusetts | 20,701 | 12,735 | 7.7 | 4.3 |
| Framingham (town) | 109 | 55 | 7.3 | 3.4 |
| Study sample | 67 | 30 | 7.3 | 3.0 |
| Ages 35-64 | | | | |
| Massachusetts, 1954-1956 | 27,554 | 16,629 | 11.0 | 5.8 |
| Framingham (town), 1953-1956 | 197 | 125 | 10.3 | 5.8 |
| Study sample, 1953-1957 | 125 | 71 | 8.9 | 4.3 |
| Respondents | 76 | 36 | 8.2 | 3.2 |
| Nonrespondents | 49 | 35 | 10.4 | 6.9 |
| Refused | 43 | 29 | 11.9 | 7.3 |
| Moved | 4 | 3 | 4.6 | 3.6 |
| Incapacitated | 2 | 3 | 14.8 | 11.9 |
| Study volunteers, 1953-1957 | 11 | 11 | 7.3 | 5.2 |

Death rates are on an annual basis per 1,000 population and are age-adjusted by the direct method using the enumerated population of Massachusetts in 1950 as the standard. Framingham rates for 1950-1952 are corrected for the institutional population.

been only the terminally ill, this group would have died during the first years of the study. With the passage of time mortality in the respondent and nonrespondent groups might be expected to converge, just as mortality in insured groups tends to rise toward the level of mortality in the general population. It need not be anticipated that mortality in the respondent group would rise completely to the level of mortality in the nonrespondent group, for if there were

substantial differences between the two groups mortality among the respondents might remain below that for nonrespondents indefinitely. However, the differential ought to become less.

It is perhaps too early to decide finally whether mortality for the two groups is or is not converging. For men, there is a suggestion that it is converging; for women the mortality differential remains very large. Unfortunately, the data for men are suspect, for there is an indication that mortality follow-up for nonrespondent men may be incomplete at this time. This would not be surprising, since the nonrespondent group includes a number of people who moved from Framingham and it is increasingly difficult to maintain complete mortality follow-up for this group.

The suggestion of incomplete mortality follow-up arises from the following observation: If the respondent and nonrespondent groups are combined, mortality during 1950-1952 is the same for the sample as for the town. During 1953-1957 this is not true; mortality in the total drawn sample is below that reported for Framingham. The amount of the apparent deficit is such that there is a real possibility that the apparent convergence of mortality in the two groups of men is mainly an artifact.

It is, of course, possible that this "deficit" comes from the respondent group rather than the nonrespondent. It is conceivable, for example, that continued medical surveillance improves the prospects of survival and that, as a consequence, mortality in the group receiving continuous clinic examination at Framingham will not rise to the level of mortality in the town at any time during the period of the study. Careful attention is paid by the clinic physicians to any cardiovascular or other abnormality appearing in the course of clinic examination, and the subject's physician is informed promptly of any suspicious findings. We have no satisfactory measure of the effect of this on the progression of disease but it must be recognized as a possible source of bias in the study.

In any event, nonresponse is a formidable problem in any study involving clinical examination of a general population. If the Framingham study had been designed to measure the prevalence or incidence of cardiovascular conditions, the low response rate would have seriously compromised our results. Even the purposes for which the study was intended, namely, the measurement of associations by means of *relative incidence*, would have been more comfortably served by complete response.

However, nonresponse exists and we believe it will continue to exist. There are cooperators and noncooperators. During the initial examination a sample of nonrespondents was intensively canvassed by the health educator after all routine procedures had failed. The additional yield was negligible. Recently a sample of nonrespondents known to be living and still resident in Framingham was again asked to come in for examination, in the hope that time had softened their objections. Again the response was negligible.

Additional insight to this problem may be gathered from a consideration of the volunteer group. At the end of the first examination round there were in the files records for a large group of volunteers not in the drawn sample. Because the number of sample respondents had fallen short of expectations, it

was decided to recall some of the volunteers as a supplement to the regular study group. The 1951 town list of residents was used as a roster. Those volunteers whose names were missing from the list were considered as moved and no further effort was made to trace them. Those on the list were asked to return if they fell within the proper age range. The resultant group included a slightly larger proportion of women than the respondent sample—57.8 as against 54.7 per cent—but was not especially different in age structure.

In line with the original study philosophy, it was at first planned to re-examine only those volunteers "normal" on their initial examination and 13 people were eliminated for hypertensive or coronary heart disease. These omissions, however, modified the clinical characteristics of the group only trivially.

This volunteer group was, as noted, drawn from people who had already been examined once at the clinic and were requested to return for re-examination. Some 3 per cent refused to do so, while 12 per cent were unavailable because they had moved out of Framingham (Table V). In contrast, 9 per cent of the sample group who received the first examination refused to take the second examination and 2 per cent were unavailable because they were no longer resident in the town.

OL.
10
959

TABLE V. RESPONSE AND NONRESPONSE BY REASON
(Framingham Study: Volunteers)

| RESPONSE AND REASON | NUMBER | PER CENT |
|---|---|---|
| Total* | 888 | 100.0 |
| Examined on Exam II | 740 | 83.3 |
| Not examined on Exam II | 148 | 16.7 |
| Moved† | 102 | 11.5 |
| Before Jan. 1, 1951 | 81 | 9.1 |
| After Jan. 1, 1951 | 21 | 2.4 |
| Died | 6 | 0.7 |
| Medical rejects‡ | 13 | 1.5 |
| Refused | 27 | 3.0 |

*Aged 30-61 as of Jan. 1, 1951. Volunteers outside this age range were not to be recalled.

†A volunteer not appearing on the town list of residents for Jan. 1, 1951, was considered to have moved from Framingham before Jan. 1, 1951, and no effort was made to trace him. If he was on the list for Jan. 1, 1951, but could not be located or was reported to have moved, he is considered as having moved after Jan. 1, 1951. There were, in addition, some volunteers not on the town list for any of the years 1949, 1950, or 1951. This group is considered nonresident.

‡Eight were omitted for hypertensive heart disease, five for coronary heart disease.

It should be noted in evaluating these figures that no effort was made to bring into the study those people who had moved out of Framingham, either from the sample or volunteer group. This was a deliberate policy to maximize follow-up, so that in a sense, part of the follow-up was bought at the expense of nonresponse. To a degree, this is also true for the Framingham residents who refused to be examined. While a few of these might have come in had greater persuasion been used, they would almost surely have returned for clinic re-examination less frequently than the present group of respondents.

FOLLOW-UP

Adequacy of follow-up can be measured only in terms of the purposes to be served. A person known to have died of a myocardial infarction one year after his first examination is successfully followed for the purposes of evaluating factors associated with the development of *de novo* coronary heart disease but his follow-up is incomplete for the purposes of evaluating the natural history of hypertension. Similarly, information from hospitals or physicians may be adequate for some purposes and not for others.

For most of the purposes of the study, follow-up of the Framingham study group has been relatively good. At the end of 6 years' follow-up there appear to be only 4 people whose mortality status cannot be specified with some assurance (Table VI) and some information is available about almost everyone who has not attended the clinic regularly (Table VII). There is only one hospital in Framingham and Framingham residents seem to make little use of hospital facilities elsewhere. There is a high certainty that any serious illness among study members still in the community would come to the clinic's attention, either from the hospital record or the report of friends or relatives in the study, since both sources of information are systematically canvassed.

VOL. 10 1959

The largest number of drop-outs from the clinic examination occurred at the time of the second examination, when 417 of the 4,469 sample respondents failed to return (Table VIII). These 417 drop-outs included 32 who died in the 2-year interval, 104 who moved, and 281 who refused re-examination at that time. However, 30 of those who moved and 116 of those who refused returned for a later examination. Thus, the apparent loss to follow-up of 9 per cent (385/4,437) of the surviving sample between the initial and second examination was not a measure of ultimate loss to clinical follow-up, nor did net loss cumulate at any such rate. In fact, at the end of 6 years (i.e., the fourth exam), the net loss to clinical follow-up of the surviving population was only 12 per cent, and a small number of those have subsequently been re-examined. Losses in the volunteer group have been nominal.

While there is a real difference in rates for clinic follow-up between the sample and volunteer group, it is not as great as may appear at first sight. The volunteer group was screened twice—on the first examination and on the second. If only those who took the second examination are considered, the loss in the subsequent 4 years because of refusal was 3.8 per cent for the sample group as against 1.8 per cent for the volunteer group, while the loss due to moving from Framingham was 2.5 per cent and 2.0 per cent, respectively. A similar contrast may be seen by comparing the first thousand examined in the sample with the last 469 (Table IX). The first thousand were in fact a group of volunteers whereas the last 469 included a sizable proportion who were persuaded to come in only with difficulty, sometimes after repeated postponements. In the former group the refusal rate for the first 6 years was 2.9 per cent; in the latter it was 13.2 per cent. The loss from move-out also rose from 1.9 to 4.7 per cent. (Since these rates are computed for 6 years' experience, they are not directly comparable with the preceding rates.)

TABLE VI. SIX-YEAR FOLLOW-UP

(Framingham Study: Respondent Sample and Volunteers)

| | RESPONDENT SAMPLE* | | VOLUNTEER | |
|---|---|---|---|---|
| | NUMBER | PER CENT | NUMBER | PER CENT |
| Total | 4,469 | 100.0 | 740 | 100.0 |
| Examined at clinic | 3,874 | 86.7 | 704 | 95.1 |
| Exam IV | 3,843 | 86.0 | 698 | 94.3 |
| Exam V,† not IV | 31 | 0.7 | 6 | 0.8 |
| Not examined at clinic | 595 | 13.3 | 36 | 4.9 |
| Died before Exam IV | 120 | 2.7 | 16 | 2.2 |
| Known living at Exam IV | 471 | 10.5 | 20 | 2.7 |
| Unknown | 4 | 0.1 | — | — |

*Includes only those found to be within the prescribed age range on Exam I. This is 25 less than the total given in Tables I and II.
†As of Aug. 6, 1958.

TABLE VII. SOME EXTRACLINIC INFORMATION

(Framingham Study: Sample Respondents and Volunteers Not Receiving Exam IV)

| INFORMATION | RESPONDENT | | |
|---|---|---|---|
| | TOTAL | SAMPLE | VOLUNTEERS |
| For deaths* | (136) | (120) | (16) |
| Autopsy | 23 | 21 | 2 |
| Report of terminal hospitalization | 27 | 21 | 6 |
| Report of attending physician | 3 | 3 | — |
| Coroner's report | 19 | 17 | 2 |
| For those not known dead† | (532) | (506) | (26) |
| Hospital record | 81 | 78 | 3 |
| Record of nonfatal coronary heart disease | 5 | 5 | — |
| Self-report | 336 | 320 | 16 |
| Working status | | | |
| Full time | 425 | 405 | 20 |
| Part time | 22 | 21 | 1 |
| Not working or retired | 37 | 34 | 3 |
| Unknown | 47 | 46 | 1 |
| Availability for examination outside clinic | | | |
| No | 277 | 267 | 10 |
| Yes | 84 | 82 | 2 |
| Maybe | 65 | 54 | 11 |
| Unknown | 106 | 103 | 3 |

Of the 668 not receiving Exam IV, 37 have returned for subsequent examination (Aug. 6, 1958). Information is for the period within 6 years of initial examination.
*Exclusive categories with priority in order listed.
†Categories not exclusive.

By inference, groups lost to clinic follow-up, when characterized according to reason for loss, may tell us something about the group of original nonrespondents. In addition, the withdrawal of some groups from clinic follow-up may

TABLE VIII. CLINIC FOLLOW-UP

(Framingham Study: Respondent Sample and Volunteers)

| TYPE OF SUBJECT AND EXAMINATION | EXAMINED | LOST SINCE PRECEDING EXAM* | | | | SKIPPED EXAM† | | |
|---|---|---|---|---|---|---|---|---|
| | | TOTAL | DEAD‡ | MOVED | REFUSED | TOTAL | MOVED | REFUSED |
| *Number* | | | | | | | | |
| Respondent sample§ | | | | | | | | |
| II | 4,052 | 271 | 32 | 74 | 165 | 146 | 30 | 116 |
| III | 3,935 | 156 | 41 | 47 | 68 | 73 | 21 | 52 |
| IV | 3,843 | 170 | 29 | 54 | 87 | 17 | 10 | 7 |
| Volunteers‖ | | | | | | | | |
| III | 718 | 14 | 4 | 3 | 7 | 8 | — | 8 |
| IV | 698 | 22 | 11 | 5 | 6 | 4 | 1 | 3 |
| *Per cent* | | | | | | | | |
| Respondent sample§ | | | | | | | | |
| II | 90.7 | 6.1 | 0.7 | 1.7 | 3.7 | 3.3 | 0.7 | 2.6 |
| III | 88.0 | 3.5 | 0.9 | 1.1 | 1.5 | 1.6 | 0.5 | 1.2 |
| IV | 86.0 | 3.8 | 0.6 | 1.2 | 1.9 | 0.4 | 0.2 | 0.2 |
| Volunteers‖ | | | | | | | | |
| III | 97.0 | 1.9 | 0.5 | 0.4 | 0.9 | 1.1 | — | 1.1 |
| IV | 94.3 | 3.0 | 1.5 | 0.7 | 0.8 | 0.5 | 0.1 | 0.4 |

*"Lost" means a person did not take Exam IV and had not taken Exam V as of Aug. 6, 1958.

†"Skipped" means a person failed to take the specified exam (and in some cases a later exam, as well) but did return on or after Exam IV.

‡Excluded from the deaths are 18 sample and 1 volunteer death counted as a "move" or "refuse" on a previous examination.

§There were 4,469 examined on Exam I. Percentages for respondent sample computed on this base.

‖There were 740 examined on Exams I and II. Percentages for volunteers were computed on this base.

Since a person is counted only the first time he misses an examination the entries need not balance.

TABLE IX. LOSS TO CLINIC FOLLOW-UP AS OF EXAM IV ACCORDING TO REASON FOR LOSS AND EXAMINATION ORDER

(Framingham Study: Respondent Sample)

| EXAMINATION ORDER | NOT EXAMINED ON OR AFTER EXAM IV | | | |
|---|---|---|---|---|
| | TOTAL | DEAD | MOVED | REFUSED |
| First thousand | 69 | 21 | 19 | 29 |
| Second thousand | 107 | 26 | 37 | 44 |
| Third thousand | 140 | 27 | 47 | 66 |
| Fourth thousand | 174 | 25 | 48 | 101 |
| Last 469* | 224 | 45 | 47 | 132 |
| (Actual) | (105) | (21) | (22) | (62) |

*Last 469 expressed as rate per 1,000.

affect the composition of the study group and modify the conclusions that can be drawn with respect to changes occurring over time for those characteristics which must be measured on the basis of clinic examination.

The loss to clinic follow-up may be ascribed to three sources: death, moving from the area, and refusal. These various groups were similar in their age-sex composition to their counterparts in the nonrespondent group (Table X). Mortality is naturally heavier among men and at older ages. The refusal rate is about the same for men and women, but, like death, refusal is more common among older subjects than among younger. On the other hand, losses due to moving from the Farmingham area are more common among young people than old and slightly more common for men than women.

TABLE X. NUMBER LOST AND AVERAGE AGE BY SEX AND REASON FOR LOSS
(Framingham Study: Drawn Sample and Respondent Sample)

| SEX AND REASON FOR LOSS | NUMBER | | PER CENT | | AVERAGE AGE* | |
|---|---|---|---|---|---|---|
| | DRAWN SAMPLE | RESPONDENT SAMPLE† | DRAWN SAMPLE | RESPONDENT SAMPLE† | DRAWN SAMPLE | RESPONDENT SAMPLE† |
| Both Sexes | 6,532 | 4,469 | —— | —— | —— | —— |
| Men | | | | | | |
| Total | 3,086 | 2,024 | 100.0 | 100.0 | 44.1 | 43.7 |
| Lost‡ | 1,050 | 294 | 34.0 | 14.5 | 44.4 | 44.4 |
| Moved | 233 | 87 | 7.6 | 4.3 | 41.2 | 39.4 |
| Died§ | 50 | 73 | 1.6 | 3.6 | 51.3 | 48.9 |
| Refused | 738 | 134 | 23.9 | 6.6 | 44.9 | 45.3 |
| Ill or incapacitated | 29 | —— | 0.9 | —— | 42.7 | —— |
| Women | | | | | | |
| Total | 3,446 | 2,445 | 100.0 | 100.0 | 44.1 | 43.6 |
| Lost‡ | 988 | 301 | 28.7 | 12.3 | 44.9 | 43.9 |
| Moved | 193 | 88 | 5.6 | 3.6 | 42.3 | 40.8 |
| Died§ | 24 | 29 | 0.7 | 1.2 | 48.0 | 48.9 |
| Refused | 726 | 184 | 21.1 | 7.5 | 45.2 | 44.7 |
| Ill or incapacitated | 45 | —— | 1.3 | —— | 48.2 | —— |

*Age is that given on the town List.

†Includes only those found to be within the specified age range at Exam I. This is 25 less than the totals given in Tables I and II.

‡Losses from the drawn sample are those who never entered the study. Losses from the respondent sample are those not receiving Exam IV or V (to Aug. 6, 1958.)

§If a person in the respondent sample was lost to follow-up because of refusing or moving and then died he is counted here as refusing or moving, not died.

We have evaluated these three groups lost to clinic follow-up in terms of a few physiologic and clinical characteristics as defined at entry to the study (Table XI). The men who died were distinguished from those who did not by higher blood pressures and a higher prevalence of cardiovascular disease and x-ray and electrocardiographic abnormalities. For most comparisons, these contrasts were even more evident for women. In addition, women had an elevated serum cholesterol and there was a small (but statistically not significant) elevation of relative weight.

TABLE XI. SELECTED CHARACTERISTICS OF SOME GROUPS NOT EXAMINED ON OR AFTER EXAM IV
(Framingham Study: Sample Respondents)

| CHARACTERISTIC | REFUSED EXAM II, NEVER RETURNED | | | MOVED EXAM II, NEVER RETURNED | | | DIED BEFORE EXAM IV* | | |
|---|---|---|---|---|---|---|---|---|---|
| | NUMBER | MEAN | | NUMBER | MEAN | | NUMBER | MEAN | |
| | | ACUTAL | EXPECTED | | ACTUAL | EXPECTED | | ACTUAL | EXPECTED |
| *Some measured characteristics* | | | | | | | | | |
| Men | | | | | | | | | |
| Serum cholesterol | 57 | 214.7 | 222.6 | 29 | 199.4 | 219.8 | 63 | 225.4 | 223.6 |
| Systolic blood pressure | 67 | 142.8 | 139.2 | 36 | 129.4 | 134.9 | 83 | 153.6 | 141.5 |
| Diastolic blood pressure | 67 | 90.4 | 87.1 | 36 | 81.8 | 85.7 | 83 | 93.2 | 87.6 |
| Relative weight | 67 | 103.3 | 101.6 | 35 | 97.8 | 101.0 | 81 | 100.4 | 101.8 |
| Women | | | | | | | | | |
| Serum cholesterol | 68 | 235.7 | 224.9 | 24 | 209.0 | 210.9 | 24 | 250.0 | 232.5 |
| Systolic blood pressure | 75 | 148.5 | 139.8 | 32 | 129.8 | 130.9 | 37 | 164.9 | 145.1 |
| Diastolic blood pressure | 75 | 88.0 | 86.0 | 32 | 81.1 | 82.1 | 37 | 96.5 | 87.9 |
| Relative weight | 75 | 105.8 | 103.6 | 32 | 100.0 | 100.0 | 37 | 107.9 | 105.8 |
| *Some counted characteristics (Number)* | | | | | | | | | |
| Men | | | | | | | | | |
| With cardiovascular disease | | 51 | 40.5 | | 11 | 21.8 | | 68 | 50.2 |
| With cardiovascular abnormality | | 34 | 27.0 | | 8 | 14.5 | | 59 | 33.4 |
| Women | | | | | | | | | |
| With cardiovascular disease | | 53 | 38.4 | | 11 | 16.4 | | 31 | 19.0 |
| With cardiovascular abnormality | | 34 | 27.4 | | 8 | 11.7 | | 22 | 13.5 |

*Includes some deaths already "lost" because of refusal or move.

"Cardiovascular disease" includes hypertension, heart disease, or cerebrovascular disease.

"Cardiovascular abnormality" includes x-ray or electrocardiographic abnormalities of the heart or peripheral vascular disease. All characteristics measured at entry.

VOL. 10, NO. 3, 1959

The second group accounting for incomplete clinic follow-up are those people still living in Framingham who did not come in for their re-examination when requested. We may describe this group as "refusals" although repeated postponement is much more common than outright refusal. There is a suggestion that clinical and physiologic abnormalities are more common among men who refused re-examination than among men who did not, but it is not a very strong suggestion. The suggestion was much stronger for women, however; in contrast with what was found for men, cholesterol and blood pressure levels were elevated in the group of women who refused to be re-examined although not so elevated as for women who died.

Again, these data suggest something about the original nonrespondent group. They give roughly the same impression as the mortality data, which is that the nonrespondent group includes a slightly higher prevalence of clinical and physiologic abnormalities than the respondent. What is more, the data point to a distinctly greater disparity for women than for men.

The last group to consider is that made up of those who missed examinations because they moved from the Framingham area. Of those missing Exam IV for this reason, two-thirds resided outside Massachusetts. Most people who have moved to other places within the state continue to return for re-examination, at least intermittently, as do a surprising number now residing outside the state. Numerically, this is the smallest of the three groups not followed at the clinic. The women in this group seem to be similar to the study group as a whole in clinical status. The men, however, seem to have lower levels of cholesterol and blood pressure and fewer cardiovascular abnormalities or disease. The picture is substantially the same no matter whether the move occurred shortly after the initial examination or after several years' follow-up or whether the move led to intermittent follow-up or complete loss to follow-up. On the other hand, it is possible that men with cardiovascular abnormalities at entry are more likely to take a long trip to the clinic and so are more likely to return regularly for examination even if they move from Framingham. On the present evidence, however, it does seem that men in the study group who move from Framingham are slightly healthier, on the average, than men who do not; whereas for women there is no clear health differential.

This may permit some inferences about the original nonrespondent group. Of the 2,038 in that group, 421 had moved out of Framingham during the first round of examinations. The follow-up data would suggest that this group of movers is not markedly different in health status from the respondent group, although the men who move seem to be slightly healthier.

## THE STUDY SCHEME: EVALUATION

In the light of this experience, what can be said of the Framingham study scheme?

In the first place, it is clear now that community studies of chronic disease which require clinic visits are not likely to have a high level of participation. Experience with other studies in the last decade indicates that the level of response at Framingham was, in fact, what might be expected.[5-7]

On the other hand, when the response rate is only two in three, can the study group be considered to differ from a set of volunteers? The answer depends upon what is meant by "volunteer." No one could, of course, enforce attendance at the clinic, but it is perfectly clear that some people came to the clinic reluctantly under some form of community pressure. Moreover, some of the people who came willingly enough would not have come if they had not been personally invited.

As we have indicated, it is reasonably certain that the healthy are over-represented in the respondent group and the ill are under-represented. To some extent we can control for this bias by our method of analysis, restricting evaluation to the cohorts classified as "normal" at the time of initial examination and measuring the characteristics under study *prior* to the development of disease. In analyzing factors associated with the development of coronary heart disease, for example, we have considered only the cohorts free of clinically recognizable coronary heart disease at entry.[3] It is our conviction that this approach sharply reduces the possibilities of bias, although the risks must always be specifically considered. The likelihood of a person's having foreknowledge of coronary heart disease is, on the present evidence, very small, although there *may* be some factors, say, an intermittent unacknowledged chest pain or a family history of coronary heart disease, which influence a person's willingness to appear at the clinic and *may* be related to his risk of subsequently developing manifest disease. What is more, many of the characteristics which are being studied are unlikely to be known to the participant; most blood chemistry would come under that rule (or at least did during the period when the initial examinations were being carried out).

The effect of bias in initial response can be further reduced by directing analysis to highly specific groups—insofar as the size of the study group allows, and this is admittedly a severe restriction. Thus, if people with high blood pressure are more likely to come into the clinic than people with low blood pressure, this per se becomes a matter of indifference so long as the analysis is made specific for each blood pressure level—unless blood pressure is highly correlated with some other critical variables which are not allowed for in the analysis. (Age and sex are two good examples of variables which can be allowed for in analysis.) If every characteristic related to the development of coronary heart disease were measured, with perhaps a few of them replaced by a closely related set of characteristics, and if the population were accurately characterized with respect to each of them, the differentials associated with each should be delineated by the specific incidence figures, even in a biased sample. It might be added that an unbiased sample is of little comfort unless these analytical conditions can be met.

These arguments were sufficiently attractive that when faced with a deficit in the study group a supplement was drawn from a group of volunteers and requested to continue taking clinic examinations. This group has since been followed exactly as the sample group has been and data from the volunteers have been used to supplement the findings in the sample group.

In the face of these numerous difficulties it may be asked whether we still consider it desirable to draw a sample of a general population group for studies

of this type. Special groups can be found which permit relatively complete coverage and follow-up. Some industrial groups would fall in that category. Volunteers can be obtained who will consent to much more intensive surveillance than it has been considered prudent to attempt at Framingham, for fear of alienating the study group from clinic follow-up.

The answer, we think, lies in the needs of inference. It is a fundamental thesis of modern experimental statistics as propounded by Fisher[10] that valid inference as to causal relations is possible only if the characteristics under study (say blood pressure or serum cholesterol) are assigned at random to the study group. Clearly, neither the Framingham study nor any other survey could be designed on that basis, since the characteristics being studied had to be accepted as found, even where experimental manipulation was conceivable. However, by studying a general population rather than a completely self-selected group diversification is introduced, a quasi-randomization, so to speak, which tends to increase our confidence in the general applicability of the study findings, particularly as they relate to associations of factors observed before overt disease occurs.

OL.
10
959

This may prove illusory. When we find, as has been found in the Framingham population, that there are differentials in the incidence of new coronary heart disease associated with differences in serum cholesterol or blood pressure levels, we must acknowledge that a number of inapparent artifacts are conceivable. One possibility is that neither of these factors would be associated with the incidence of disease if some other factor, X, were not characteristic of this population. Factor X could be some environmental factor such as the presence or absence of a trace element or adulterant in the food or water supply, or an air contaminant, or alternatively the presence or absence of some genetic characteristic. We know of no way to guarantee against this threat other than repetition of the study at other times and in other places (which is being done), just as laboratory experiments must be repeated under varying conditions. The community study does, however, provide some insurance in that a wide variety of occupational situations, familial situations, and patterns of energy expenditure are included in the total environment of the group.

Another potential source of difficulty arises from the fact that the range of measured variables under examination at Framingham, such as blood pressure or cholesterol, does not include the full human range of variability. We know of no way to guard against the possibility that we are in fact studying a peculiar group with respect to these variables except to restrict the inferences that are made to the range of data observed in this study. It is a plausible hypothesis, for example, that there exists a threshold level of serum cholesterol below which there is hardly any risk of coronary disease, and that below that level there is no independent association of blood pressure with coronary disease. This is not suggested by data from the Framingham study group, but it may be true for some populations living in Japan (to consider one possibility) with mean cholesterol levels substantially below all but the lowest levels found in comparable age-sex groups in Framingham.

Another hazard is what we may call the silent partner. We can conceive of such a situation in the present problem. *If* anxiety is in some fashion associated with hypertension (whether as cause, consequence, or merely as concommitant variation), and *if*, as would be a reasonable hypothesis, anxiety is associated, either positively or negatively, with willingness to join the study group, and *if* anxiety itself makes some independent contributions to coronary risk, *then* any one of several situations might conceivably occur. Two of them are: (1) if hypertension and anxiety are positively correlated, and increased anxiety is associated with decreased response, then within each age-sex group subclassified by hypertension the gradient of coronary risk in relation to hypertension would tend to be decreased; and (2) if, on the other hand, anxiety increased response, it would be reasonable to assume that there would be proportional over-response of those who were both anxious and hypertensive. In this instance the gradient of coronary risk in relation to hypertension would tend to be increased.

Granted that these and other hazards do exist in a community study, it is also true that such a study has an important and in some respects a unique contribution to the understanding of cardiovascular disease. There is sometimes a tendency to regard an experiment as providing a higher form of evidence. This seems to us completely unrealistic. Certainly the classical experiment is a powerful tool for investigation, but it can be neither designed nor interpreted except by reference to observed unmanipulated reality. The community study provides this kind of reference for the investigation of chronic diseases in human populations. What is more, the simplified model of reality on which an experiment is based may be so inappropriate or the conditions under which the experiment is carried out may be so peculiar that the results are in fact useless. Thus, the means used to manipulate serum cholesterol levels by dietary experiments seem to bear little resemblance to the mechanisms that in fact control serum cholesterol levels in human populations and may ultimately provide little insight into the factors primarily responsible for these levels.

In the same connection, we do not agree with the thesis, if anyone seriously holds such a position, that only experiments allow of reasonable probability statements. The conditions under which a specific probability statement is admissable are not our immediate concern, but we have no doubt in our own minds that a probability statement may be as legitimate and useful for inquiries based on observational studies of human populations as for classical experiments.

On the other hand a statistical statement must begin and must end with the probability model. Thus, when you compare the mean blood pressures for a coronary and a noncoronary group, the probability statement you make refers to two sets of numbers assumed to conform to a specified probability behavior, not to blood pressure or coronary disease. You may decide that the means of the two sets of numbers are or are not the same, by whatever statistical rules you have chosen to use. This is a statistical decision. When you go beyond this to make assertions about coronary heart disease and blood pressure, the inference may or may not be justified, but it is not statistical.

OTHER STUDY SCHEMES: RETROSPECTIVE STUDIES

The value of the prospective study scheme at Framingham does not, by any means, rule out the utility of other types of epidemiologic studies. In particular, retrospective studies are often very useful devices for obtaining information quickly and economically. They do, however, include serious hazards of selective bias, and on occasion it is difficult to sort events recorded retrospectively into their proper sequence for the purpose of inferring cause and effect.[11]

The Framingham experience does not bear directly on many of the problems of retrospective studies, but it does throw some light on the effect of the coronary event itself on the characteristics under study. This has been a source of concern in retrospective studies of coronary heart disease. Thus, blood pressure may drop after the coronary event or be reduced by clinical management. The event may lead to weight reduction or to changes in diet or activity that affect serum cholesterol levels. Death may be a selective factor; a person may be more likely to die in his first coronary attack if his blood pressure is high, or he is obese, or his serum cholesterol level is greatly elevated.

OL.
10
959

Generally speaking, these concerns are not borne out by the Framingham data. There is no suggestion that serum cholesterol levels are different after recovery from the coronary event than the levels recorded in advance of the event (Table XII). Weight decreases slightly for men but not for women. While blood pressure decreases for both men and women, this is largely, if not entirely, explained by a tendency of blood pressures to regress in the study group as a whole, possibly because the examination situation became more familiar and hence less threatening. There is some suggestion, especially for men, that abrupt shifts from definite hypertension to normotension are more common

TABLE XII. MEAN LEVELS OF CHOLESTEROL, FRAMINGHAM RELATIVE WEIGHT, AND BLOOD PRESSURE ON THE EXAMINATION JUST BEFORE AND JUST AFTER CORONARY HEART DISEASE EVENT

(Framingham Study: New Coronary Heart Disease in Six Years' Follow-up)

| SEX AND EXAMINATION | CHOLESTEROL | | FRAMINGHAM RELATIVE WEIGHT | | BLOOD PRESSURE | | |
|---|---|---|---|---|---|---|---|
| | NUMBER MEASURED | MEAN | NUMBER MEASURED | MEAN | NUMBER MEASURED | MEAN | |
| | | | | | | SYSTOLIC | DIASTOLIC |
| Men | | | | | | | |
| Exam before | 84 | 243.5 | 96 | 104.5 | 99 | 148.7 | 91.1 |
| Exam after | 84 | 246.3 | 96 | 101.8 | 99 | 141.4 | 87.4 |
| Women | | | | | | | |
| Exam before | 33 | 276.8 | 50 | 111.7 | 50 | 163.7 | 96.0 |
| Exam after | 33 | 278.8 | 50 | 110.5 | 50 | 162.3 | 94.7 |

Data are for respondent sample and volunteers free of coronary heart disease at entry.

after an event than could be accounted for by the usual permutations of blood pressure (Table XIII), but for some reason the net effect on mean blood pressures does not reflect this.

TABLE XIII. CHANGES IN DIAGNOSIS OF DEFINITE HYPERTENSION
(Framingham Study: Population at Risk and New Coronary Heart Disease in Six Years' Follow-up)

| SEX AND HYPERTENSION DIAGNOSIS ON EXAM AFTER EVENT | POPULATION AT RISK* | NEW CORONARY HEART DISEASE† |
|---|---|---|
| Men | | |
| Total | 290 | 33 |
| Definite | 170 | 21 |
| Borderline | 96 | 4 |
| Normotensive | 24 | 8 |
| Women | | |
| Total | 389 | 22 |
| Definite | 278 | 19 |
| Borderline | 89 | 3 |
| Normotensive | 22 | — |

*Those free of coronary heart disease at entry who were diagnosed as definitely hypertensive on Exam II. Distributed according to hypertension diagnosis on Exam III.

†Those persons diagnosed as definitely hypertensive on examination before the first appearance of coronary heart disease. Distributed according to hypertension diagnosis on examination after the event.

Naturally, not all characteristics will remain unaffected by the event itself and, in particular, attitudes and behavior may be considerably altered by a serious illness. Furthermore, characteristics which cannot be measured objectively are more suspect than those that can. However, serum cholesterol levels, blood pressure, and weight are important variables for the study of coronary heart disease and it is of great interest that they are apparently the same after recovery from the event as before the event.

What is more, relative weight, blood pressure and cholesterol levels for men who develop new coronary heart disease are the same for those who die from the disease as for those who do not (Table XIV). This is not the case for women, however. Those women who died weighed less, but had higher cholesterol levels, than those who survived to Exam IV. However, death was relatively infrequent among women, accounting for only 8 of the 61 new events, so that its effect as a selective process would not be very marked, under any circumstances.

For men, on the other hand, those who die shortly after the first clinical manifestation of coronary heart disease constitute a sizable proportion of those who develop the disease; so that it is of considerable moment to know that in some important respects they appear to be similar to those who develop the disease and survive, that is to say, those available for retrospective study in the clinic. It is not contended that death is associated with no differentials of interest in the investigation of coronary heart disease epidemiology. We know that it is associated with smoking differentials among men, and it may be asso-

VOL 10 195

OL.
10
959



TABLE XIV. MEAN LEVELS OF SERUM CHOLESTEROL, BLOOD PRESSURE, AND RELATIVE WEIGHT AT ENTRY
(Framingham Study: Survivors and Decedents of New Coronary Heart Disease in Six Years' Follow-up)

| AGE AND SEX | SERUM CHOLESTEROL | | | | BLOOD PRESSURE | | | | | | FRAMINGHAM RELATIVE WEIGHT | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | NUMBER | | MEAN | | NUMBER | | MEAN SYSTOLIC | | MEAN DIASTOLIC | | NUMBER | | MEAN | |
| | LIVING | DEAD | LIVING | DEAD | LIVING | DEAD | LIVING | DEAD | LIVING | DEAD | LIVING | DEAD | LIVING | DEAD |
| Men | 93 | 29 | 249.5 | 248.9 | 95 | 30 | 147.9 | 153.3 | 91.8 | 92.9 | 95 | 30 | 103.2 | 104.8 |
| 35-39 | 2 | 1 | 273.5 | 250.0 | 2 | 1 | 128.0 | 136.0 | 82.5 | 94.0 | 2 | 1 | 102.0 | 122.0 |
| 40-44 | 7 | 2 | 263.1 | 271.0 | 7 | 2 | 145.3 | 132.0 | 94.7 | 83.0 | 7 | 2 | 104.3 | 111.5 |
| 45-49 | 11 | 3 | 262.4 | 275.0 | 12 | 3 | 136.8 | 124.7 | 91.6 | 90.0 | 12 | 3 | 100.2 | 94.0 |
| 50-54 | 15 | 3 | 250.1 | 236.3 | 16 | 3 | 136.3 | 166.3 | 89.3 | 105.3 | 16 | 3 | 103.9 | 108.3 |
| 55-59 | 11 | 5 | 238.8 | 243.8 | 11 | 5 | 137.5 | 155.0 | 90.1 | 90.4 | 11 | 5 | 104.8 | 107.6 |
| 60-64 | 31 | 12 | 242.4 | 246.0 | 31 | 12 | 158.4 | 157.3 | 92.1 | 94.9 | 31 | 12 | 101.8 | 104.8 |
| 65-69 | 16 | 3 | 252.2 | 240.7 | 16 | 4 | 158.2 | 166.3 | 95.3 | 87.8 | 16 | 4 | 106.2 | 99.5 |
| Women | 51 | 7 | 261.5 | 285.7 | 53 | 8 | 167.2 | 168.0 | 96.6 | 96.5 | 53 | 8 | 112.3 | 100.6 |
| 40-44 | 1 | — | 479.0 | ——— | 1 | — | 112.0 | ——— | 70.0 | ——— | 1 | — | 89.0 | ——— |
| 45-49 | 2 | — | 235.0 | ——— | 2 | — | 131.0 | ——— | 86.0 | ——— | 2 | — | 95.0 | ——— |
| 50-54 | 9 | 2 | 298.3 | 235.5 | 9 | 3 | 158.9 | 150.0 | 98.8 | 90.7 | 9 | 3 | 116.3 | 82.7 |
| 55-59 | 12 | 1 | 250.4 | 292.0 | 14 | 1 | 160.4 | 270.0 | 98.0 | 135.0 | 14 | 1 | 113.5 | 152.0 |
| 60-64 | 21 | 1 | 254.8 | 400.0 | 21 | 1 | 179.1 | 124.0 | 95.8 | 90.0 | 21 | 1 | 111.2 | 113.0 |
| 65-69 | 6 | 3 | 224.5 | 279.0 | 6 | 3 | 175.0 | 166.7 | 100.7 | 91.7 | 6 | 3 | 117.2 | 97.3 |

Age as of Exam IV, estimated for decedents. Survival status also as of Exam IV.

ciated with some other differentials. But it is of great interest that such important variables as relative weight, cholesterol, and blood pressure are the same for those men who die shortly after the development of coronary heart disease as for those men who do not.

These findings, both with respect to survivors and decedents, are certainly encouraging to anyone considering a retrospective study of coronary heart disease. Naturally other factors must be weighed into the balance as well, before a retrospective study in this area is undertaken. These factors, however, are outside the scope of the present discussion.

SUMMARY

In diseases which develop slowly, such as atherosclerotic and hypertensive heart disease, the approach which is often considered most desirable is long-term surveillance of closed populations. However, there are problems in the choice of populations for study and questions about the inferences that may be drawn from a survey approach, and these issues are discussed.

The 10 years' experience from the Framingham Heart Epidemiology Study is reviewed in terms of biases in initial response and biases in follow-up by clinical and other means. While an average of 68.8 per cent of the drawn sample came to the study clinic, the response rate differed by age, sex, and according to section of town. It also appears to have differed according to certain clinical characteristics. This is inferred from a persistently higher mortality in the nonrespondent group than the respondent. It is also inferred from differences between those respondents who have been re-examined regularly and those lost to clinic follow-up. These biases are of moment in inferring from relationships shown by a description of the population (retrospective inferences) but are considered trivial for prospective evaluation of hypotheses.

It is shown that some important characteristics studied in the epidemiology of coronary heart disease are not altered permanently by the disease event.

REFERENCES

1. Dawber, T. R., Meadors, G. F., and Moore, F. E.: Epidemiological Approaches to Heart Disease: The Framingham Study, Am. J. Pub. Health **41**:279, 1951.
2. Dawber, T. R. and Moore, F. E.: Longitudinal Study of Heart Disease in Framingham, Massachusetts: An Interim Report in Research in Public Health, Annual Conference of The Milbank Memorial Fund, New York, 1952, The Milbank Memorial Fund, p. 241.
3. Dawber, T. R., Moore, F. E., and Mann, G. V.: Coronary Heart Disease in the Framingham Study, Am. J. Pub. Health **47**(no. 4, pt. 2):4, 1957.
4. Health Statistics from the U. S. National Health Survey, Series B-4: 7, U. S. D.H.E.W., Washington, D.C., June, 1958.
5. Cobb, S., King, S. and Chew, E.: Differences Between Respondents and Nonrespondents in a Morbidity Survey Involving Clinical Examination, J. CHRON. DIS. **5**:636, 1957.
6. Krueger, D. E.: Measurement of Prevalence of Chronic Disease by Household Interviews and Clinical Evaluations, Am. J. Pub. Health **47**:953, 1957.
7. Trussell, R. E.: Comparison of Various Methods of Estimation of the Prevalence of Chronic Disease in a Community. The Hunterdon County Study, Am. J. Pub. Health **46**:173, 1956.
8. Mainland, D.: Safety in Numbers, Circulation **16**:784, 1957.
9. Berkson, J.: The Statistical Study of Association Between Smoking and Lung Cancer, Proc. Staff. Meet. Mayo Clin. **30**:319, 1955.
10. Fisher, R. A.: The Design of Experiments, London, 1935, Oliver & Boyd, Ltd.
11. Dorn, H. F.: Some Applications of Biometry in the Collection and Evaluation of Medical Data, J. CHRON. DIS. **1**:638, 1955.

# SOME EPIDEMIOLOGIC ASPECTS OF THE HIGH MORTALITY RATE IN THE YOUNG WIDOWED GROUP

ARTHUR S. KRAUS, SC.D., AND
ABRAHAM M. LILIENFELD, M.D., M.P.H.,
BALTIMORE, MD.

(Received for publication April 16, 1959)

OL.
10
959

IN 1956, the National Office of Vital Statistics published comprehensive statistical data on mortality by marital status.[1] These were based on all deaths in the continental United States in 1949-1951, and on the 1950 Census, which reported the population according to marital status and other characteristics. The number of deaths and death rates for all causes and for selected causes were presented, by marital status, age, color, and sex.

These data were discussed in a brief review appearing in the *Statistical Bulletin of the Metropolitan Life Insurance Company*.[2] Some of the major points discussed were:

1. Married people experienced a lower mortality rate from all causes than did single persons, the widowed, and the divorced, for every specific age group in each sex and each color. For both sexes and colors combined, the ratios of the age-adjusted death rates in these three not-married groups to the age-adjusted rate in the married group were 1.47, 1.46, and 1.84, respectively. (The age-adjusted death rates are based on the age-specific rates for age groups between 20 and 74 for each marital status, and the age distribution of the total U. S. population between 20 and 74 in 1950.)

2. The relative excess mortality in the not-married categories compared to the married group was greater at the younger ages.

3. The relative excess mortality in the not-married categories was consistently greater in males than in females.

4. The relative excess mortality in the not-married categories was particularly high for tuberculosis in each sex, and for influenza and pneumonia, cirrhosis of the liver, suicide, and accidents (other than motor vehicle) in males. (For these causes, the age-adjusted rate in each of the three not-married groups was more than double the rate in the married group in the table appearing in the

Dr. Kraus is Chief, Division of Vital Records and Statistics, State of Maryland Department of Health. Dr. Lilienfeld is Professor of Public Health Administration and Director, Division of Chronic Diseases, The Johns Hopkins University School of Hygiene and Public Health.

Metropolitan Life Insurance Company article.) In males, leukemia and aleukemia, and "lymphosarcoma, etc." were the only causes among twenty-two leading causes shown that were not higher in every not-married category than in the married group. These two causes were slightly lower in both the single and the widowed than in the married. In females, six of the causes shown were not higher in every not-married category than in the married group, with the diabetes rate in single women only 56 per cent of the rate in married women.

### THE HIGH MORTALITY RATE AMONG THE YOUNG WIDOWED

An outstanding feature of the N.O.V.S. data,[1] only very briefly mentioned in the Metropolitan article,[2] is the great relative excess in mortality among the young widowed group, compared to the young married group.

Table I shows the death rates from all causes, by marital status, age, color, and sex, which appeared in the N.O.V.S. report.[1] The open-ended age groups, under 20 and 75 and over, are excluded, because of important differences in the age distribution within these groups between the different marital statuses. Table II presents the ratios of the death rates for the not-married groups in Table I to the rate for the married group.

TABLE I. AVERAGE ANNUAL DEATH RATES, BY MARITAL STATUS, AGE, COLOR, AND SEX, 20 TO 74-YEAR AGE GROUP, UNITED STATES, 1949-1951
(Death rates are per 100,000 population in each specified group)

| COLOR, SEX, AND MARITAL STATUS | AGE | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 20–24 | 25–34 | 35–44 | 45–54 | 55–59 | 60–64 | 65–69 | 70–74 |
| White male | | | | | | | | |
| Married | 128 | 147 | 322 | 868 | 1,716 | 2,529 | 3,653 | 5,457 |
| Single | 196 | 309 | 736 | 1,629 | 2,851 | 3,988 | 5,410 | 7,805 |
| Widowed | 496 | 633 | 928 | 1,734 | 2,631 | 3,621 | 4,851 | 6,818 |
| Divorced | 333 | 533 | 1,140 | 2,246 | 3,562 | 4,796 | 6,604 | 9,223 |
| White female | | | | | | | | |
| Married | 71 | 100 | 215 | 511 | 961 | 1,506 | 2,329 | 3,879 |
| Single | 91 | 175 | 320 | 623 | 1,088 | 1,588 | 2,435 | 4,187 |
| Widowed | 263 | 266 | 395 | 728 | 1,187 | 1,810 | 2,697 | 4,406 |
| Divorced | 143 | 223 | 391 | 718 | 1,285 | 1,998 | 3,284 | 5,773 |
| Nonwhite male | | | | | | | | |
| Married | 281 | 372 | 674 | 1,494 | 2,612 | 3,337 | 3,646 | 5,043 |
| Single | 418 | 818 | 1,813 | 3,362 | 5,428 | 6,176 | 7,105 | 9,405 |
| Widowed | 725 | 1,444 | 1,986 | 3,733 | 5,490 | 6,451 | 6,053 | 7,797 |
| Divorced | 418 | 972 | 1,545 | 3,042 | 4,738 | 5,896 | 6,714 | 8,616 |
| Nonwhite female | | | | | | | | |
| Married | 222 | 328 | 638 | 1,258 | 2,003 | 2,630 | 2,626 | 4,262 |
| Single | 315 | 601 | 1,185 | 2,018 | 2,767 | 3,917 | 3,525 | 5,651 |
| Widowed | 506 | 776 | 1,230 | 2,292 | 3,526 | 4,434 | 3,634 | 5,295 |
| Divorced | 326 | 490 | 882 | 1,678 | 2,728 | 3,950 | 3,622 | 6,574 |

It is seen in Table II that the highest ratios in any not-married group or age group are for the widowed between 20 and 34. Each of the eight ratios for these widowed exceeds 2.0, with a high of 4.31 for the white male widowed in

the 25-34 age group. Numerically, these deaths in the widowed under 35 are far from an insignificant problem. An average of 1,105 such deaths occurred annually in the United States in the 1949-1951 period, among a population of 233,000. Another 3,672 deaths occurred annually among the widowed of age 35-44, a group which still shows substantial ratios in Table II.

TABLE II. RATIOS OF DEATH RATES FOR NOT-MARRIED GROUPS IN TABLE I TO RATE FOR THE MARRIED GROUP

| COLOR, SEX, AND MARITAL STATUS | AGE | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 20–24 | 25–34 | 35–44 | 45–54 | 55–59 | 60–64 | 65–69 | 70–74 |
| White male | | | | | | | | |
| Single | 1.53 | 2.10 | 2.28 | 1.88 | 1.66 | 1.58 | 1.48 | 1.43 |
| Widowed | 3.88 | 4.31 | 2.88 | 2.00 | 1.53 | 1.43 | 1.33 | 1.25 |
| Divorced | 2.60 | 3.62 | 3.54 | 2.59 | 2.08 | 1.90 | 1.81 | 1.69 |
| White female | | | | | | | | |
| Single | 1.28 | 1.75 | 1.49 | 1.22 | 1.13 | 1.05 | 1.04 | 1.08 |
| Widowed | 3.70 | 2.66 | 1.84 | 1.42 | 1.24 | 1.20 | 1.16 | 1.14 |
| Divorced | 2.01 | 2.23 | 1.82 | 1.40 | 1.34 | 1.33 | 1.41 | 1.49 |
| Nonwhite male | | | | | | | | |
| Single | 1.49 | 2.20 | 2.69 | 2.25 | 2.08 | 1.85 | 1.95 | 1.86 |
| Widowed | 2.58 | 3.88 | 2.95 | 2.50 | 2.10 | 1.93 | 1.66 | 1.55 |
| Divorced | 1.49 | 2.61 | 2.29 | 2.04 | 1.81 | 1.77 | 1.84 | 1.71 |
| Nonwhite female | | | | | | | | |
| Single | 1.42 | 1.83 | 1.86 | 1.60 | 1.38 | 1.49 | 1.34 | 1.32 |
| Widowed | 2.28 | 2.37 | 1.93 | 1.82 | 1.76 | 1.68 | 1.38 | 1.24 |
| Divorced | 1.47 | 1.49 | 1.38 | 1.33 | 1.36 | 1.50 | 1.38 | 1.54 |

In each of the four color-sex groups, the ratios for the widowed decrease steadily with increasing age, and become lower than those for the single or the divorced at the oldest ages shown. The ratios for the male widowed are consistently higher than for the female widowed. The ratios for the white widowed are higher than for the nonwhite widowed under age 35, but are substantially lower than for the nonwhite widowed above that age.

Table III shows the death rates for twenty leading causes of death in the married and in the widowed, for three age groups between 20 and 44, by sex. Those causes are included for a sex that: (1) had a crude annual death rate for that sex in excess of 10.0 per 100,000 in the N.O.V.S. report,[1] except for causes of a miscellaneous nature such as "other disease of the heart," and, (2) had five or more deaths in the 1949-1951 period for that sex in each of the age groups 25-34 and 35-44 in each of the widowed and married groups. The ratios of the death rates in the widowed to those in the married are shown in the columns at the right of the table.

All of these ratios in Table III exceed one. However, the ratios vary a great deal for the different causes of death. The highest ratios in each cause are in the 20-24 or 25-34 age group, with substantially lower ratios in the 35-44 age

TABLE III. AVERAGE ANNUAL DEATH RATES FOR SELECTED CAUSES IN THE MARRIED AND THE WIDOWED, FOR THE 20 TO 44 AGE GROUP, BY SEX AND AGE, UNITED STATES 1949-1951
(Death Rates are per 100,000 population in each specified group)

| CAUSE* | SEX | RATES IN THE MARRIED, BY AGE | | | RATES IN THE WIDOWED, BY AGE | | | WIDOWED RATE DIVIDED BY MARRIED RATE, BY AGE | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 20–24 | 25–34 | 35–44 | 20–24 | 25–34 | 35–44 | 20–24 | 24–34 | 35–44 |
| Tuberculosis (001–019) | M | 7.3 | 11.2 | 21.8 | 58.9 | 141.8 | 168.0 | 8.1 | 12.7 | 7.7 |
| | F | 13.3 | 15.4 | 15.0 | 75.2 | 76.1 | 50.6 | 5.7 | 4.9 | 3.4 |
| Malignant neoplasms—total (140-205) | M | 9.4 | 16.2 | 43.0 | † | 37.5 | 90.3 | —— | 2.3 | 2.1 |
| | F | 7.7 | 20.9 | 73.7 | 18.5 | 52.4 | 133.7 | 2.4 | 2.5 | 1.8 |
| Of digestive organs and peritoneum (150-156A, 157-159) | M | 1.0 | 3.2 | 13.8 | —— | 12.8 | 32.7 | —— | 4.0 | 2.4 |
| | F | 0.8 | 3.2 | 13.1 | † | 9.1 | 24.5 | —— | 2.8 | 1.9 |
| Of respiratory system (160-164) | M | 0.4 | 1.1 | 7.4 | —— | 5.5 | 16.5 | —— | 5.0 | 2.2 |
| Of breast (170) | F | 0.2 | 3.7 | 19.2 | † | 7.1 | 24.8 | —— | 1.9 | 1.3 |
| Of genital organs (171-179) | F | 1.6 | 6.0 | 24.1 | 6.6 | 20.3 | 58.8 | 4.1 | 3.4 | 2.4 |
| Diabetes (260) | M | 0.7 | 1.4 | 3.1 | —— | 4.6 | 7.7 | —— | 3.3 | 2.5 |
| | F | 1.2 | 1.7 | 3.5 | † | 4.2 | 10.1 | —— | 2.5 | 2.9 |
| Vascular lesions of the central nervous system (330-334) | M | 2.2 | 3.6 | 15.8 | 25.8 | 29.3 | 63.6 | 11.7 | 8.1 | 4.0 |
| | F | 1.7 | 4.1 | 18.1 | 10.5 | 17.4 | 52.8 | 6.2 | 4.2 | 2.9 |
| Disease of the heart (410-443) | M | 5.7 | 19.0 | 105.7 | 44.2 | 97.0 | 265.6 | 7.8 | 5.1 | 2.5 |
| | F | 6.6 | 14.2 | 47.8 | 36.9 | 54.4 | 135.0 | 5.6 | 3.8 | 2.8 |
| Chronic rheumatic heart disease (410-416) | M | 2.1 | 5.5 | 12.2 | † | 10.1 | 24.2 | —— | 1.8 | 2.0 |
| | F | 3.2 | 5.8 | 12.2 | † | 10.9 | 20.0 | —— | 1.9 | 1.6 |
| Arteriosclerotic heart disease (420) | M | 1.7 | 8.6 | 71.7 | † | 42.1 | 140.9 | —— | 4.9 | 2.0 |
| | F | 1.1 | 2.8 | 14.7 | 15.8 | 16.5 | 40.4 | 14.4 | 5.9 | 2.7 |
| Nonrheumatic chronic endocarditis and other myocardial degeneration (421, 422) | M | 0.4 | 1.3 | 6.2 | † | 16.5 | 26.7 | —— | 12.7 | 4.3 |
| | F | 0.7 | 1.5 | 5.2 | 7.9 | 9.8 | 20.2 | 11.3 | 6.5 | 3.9 |
| Hypertension with heart disease (440-443) | M | 0.3 | 1.7 | 9.6 | † | 18.3 | 43.9 | —— | 10.8 | 4.6 |
| | F | 0.5 | 2.3 | 11.7 | 9.2 | 10.9 | 42.1 | 18.4 | 4.7 | 3.6 |
| General arteriosclerosis (450) | F | † | 0.0‡ | 0.2 | † | 1.1 | 1.3 | —— | 36.7 | 6.5 |
| Chronic and unspecified nephritis and other renal sclerosis (592-594) | M | 2.5 | 3.9 | 7.3 | † | 10.1 | 23.5 | —— | 2.6 | 3.2 |
| | F | 2.3 | 3.4 | 6.9 | † | 9.4 | 17.4 | —— | 2.8 | 2.5 |
| Influenza and pneumonia (480-493) | M | 1.5 | 2.6 | 7.0 | † | 20.1 | 47.1 | —— | 7.7 | 6.7 |
| | F | 2.2 | 2.7 | 5.1 | 9.2 | 13.4 | 19.9 | 4.2 | 5.0 | 3.9 |
| Cirrhosis of liver (581) | M | 0.2 | 1.3 | 8.2 | —— | 7.3 | 33.7 | —— | 5.6 | 4.1 |
| Motor vehicle accidents (E810-E835) | M | 45.5 | 32.3 | 28.0 | 139.8 | 146.4 | 88.2 | 3.1 | 4.5 | 3.2 |
| | F | 9.4 | 7.8 | 7.5 | 55.4 | 25.9 | 19.8 | 5.9 | 3.3 | 2.6 |
| All other accidents (E800-802, E840-962) | M | 35.2 | 32.3 | 35.5 | 84.6 | 99.7 | 105.1 | 2.4 | 3.1 | 3.0 |
| | F | 5.4 | 5.1 | 7.0 | 23.7 | 19.9 | 19.5 | 4.4 | 3.9 | 2.8 |
| Suicide (E963,E970-E979) | M | 7.1 | 9.5 | 17.2 | 66.2 | 65.9 | 59.7 | 9.3 | 6.9 | 3.5 |

(*For footnotes, see opposite page.*)

group, in general. For the 20-24 and 25-34 age groups, the nine ratios shown for total malignant neoplasms, diabetes, chronic rheumatic heart disease, and chronic and unspecified nephritis and other renal sclerosis are all below 3.4. In contrast, for the same age groups, the eighteen ratios for tuberculosis, vascular lesions of the central nervous system, arteriosclerotic heart disease, nonrheumatic chronic endocarditis and other myocardial degeneration, hypertension with heart disease, and general arteriosclerosis are all above 4.2, with each of these six causes having at least one ratio above 10.0. Some of the ratios for influenza and pneumonia and suicide are also substantial.

Arteriosclerotic heart disease and vascular lesions affecting the central nervous system are the first and third leading causes of death in the United States, respectively (malignant neoplasms is second). A major purpose of this report is to call attention to the markedly increased risk of dying from these two conditions among the young widowed group in comparison to the young married group, and to consider possible reasons for this phenomenon, in the hope that such discussion will stimulate more definitive types of investigation.

VOL.
10
959

## EXPLANATORY HYPOTHESES

An artificial factor that might contribute to the observations already described would involve the misstatement of marital status on death certificates and census interviews. With respect to the observed excess mortality rates in the widowed group over the married group, the misstatement of married deaths as "widowed" on death certificates, or of widowed members of the population as "married" in census interviews would be required. It seems unlikely to us that this factor plays an important role. A given degree of error in the census would produce the same degree of relative error in the rates for each cause of death in a given age-sex group, whereas Table III shows that the ratios vary greatly between different causes. For errors on the death certificates to account for the observed results, the degree of error in stating marital status would have to be much greater for certain causes than for others at a given age of death. Furthermore, whether the errors were on death certificates or census returns or both, they would have to be proportionately greater for men than for women, and would have to be much fewer relatively with advancing age. However, a study comparing stated marital status on death certificates shortly after April 1, 1950, and on 1950 Census returns for the same individuals would be useful.

Another artificial factor which might contribute to the observed picture is the probable higher average age of the widowed than of the married within each of the younger age groups shown in Table III. However, this difference would tend to be less in the narrow 20-24 age group than in the wider 35-44 group, whereas the ratios are substantially higher in the former age group. It seems unlikely that the widowed in the 20-24 group could be more than a year and a

*Numbers in parentheses after each cause are category numbers of the *Sixth Revision of the International List*, 1948.

†Rates have been omitted where the number of deaths in the 1949-1951 period for the given category was less than 5.

‡There were 10 deaths from this cause to married women aged 25-34 in 1949-1951, and approximately 10,000,000 people in the corresponding population. This gives an annual rate of 0.03 per 100,000.

half older than the married in that group on the average. Such a slight age difference should account for only a small part of relative risk factors of the magnitude observed for many causes in Table III. Furthermore, the widowed death rates in the 20-24 or 25-34 groups for a number of causes shown in Table III were even substantially higher than the married rate in the next oldest age group.

The above two hypothetical artificial factors are the only ones which have occurred to us as plausible possibilities. Since it seems unlikely to us that they account for more than a slight part of the high relative risks observed among the young widowed, we suspect that there is a genuine association between being in the young widowed population and subsequent death from the causes indicated.

Various hypotheses can be postulated for this association. One possibility is that the association is a result of "selection." This explanation assumes that the subsequent mortality rate of all those who became widowed is no different than for those of the same age who remained married. However, it hypothesizes that those who became widowed and were in good health tended to remarry and return themselves to the married population, while the ill who became widowed tended to remain widowed and be classified as such in both the census and on the death certificates if death occurred. Thus, the group remaining widowed would have higher mortality rates than the married because they tended to be "selected" in terms of ill health. It is of interest to determine if this hypothesis could account for the degree of differences noted.

Table IV shows the married and widowed population of age 14 to 24 in the United States in 1950, by sex and age.[3] According to the "selection" hypothesis, the 15,725 widowed males under age 25 in 1950 include those in whom illness has prevented remarriage. Furthermore, the other group that became widowed at the same ages and times but was in good health and had remarried is assumed to be included among the 2,391,425 married men in this age group in the Census. No doubt such a phenomenon occurs to some extent. The question is whether it happened to a great enough extent to explain the magnitude of the excess mortality observed in those who remained widowed.

TABLE IV. MARRIED AND WIDOWED POPULATION AGED 14-24, BY SEX AND AGE, UNITED STATES, 1950

| AGE | MALES | | FEMALES | |
|---|---|---|---|---|
| | MARRIED | WIDOWED | MARRIED | WIDOWED |
| Total under 25 | 2,391,425 | 15,725 | 4,751,355 | 31,105 |
| 14 | 6,660 | 1,670 | 6,980 | 565 |
| 15–19 | 166,955 | 4,995 | 887,615 | 5,260 |
| 20–24 | 2,217,810 | 9,060 | 3,856,760 | 25,280 |

For purposes of discussion, we can make the extreme assumption that an of those who became widowed at the same times and ages as the "widowed" in the 1950 Census, but had remarried before the Census, experienced no deaths in

the 1949-1951 period. Then, if the number in this group equaled the number in the group that remained widowed, the true death rate for the total group that had become widowed would be half of the death rate in the group that remained widowed. If "selection" is to explain a mortality rate in the group that remained widowed that was ten times as great as that in the married group, the number who had become widowed and remarried would have to be approximately nine times the number who remained widowed. (The married group is so large that the addition to it of a healthy group nine times the size of the group that remained widowed would produce a very slight reduction in death rates for the married. Almost the whole effect of such remarriage would be to multiply the death rates for those who remained widowed.)

A report by the National Office of Vital Statistics[4] gives statistics for persons who married between January, 1947, and June, 1954, a 7½-year period. This report shows that 14,000 widowed males had remarried at ages 12-24 during the period (the ages were not shown in greater detail). This would be an average of 1,867 per year. Assuming that these were distributed by age as were those who remained widowed, there were 1,076 remarriages annually at ages 20-24 among previously widowed males.

It is not known for how long the widowed in the 1950 Census had been widowed. However, an issue of *Current Population Reports*[5] provides some data on duration of present marital status for those included in the April, 1948, monthly population sample survey. This shows that widowed males in the 14-34 age group had been widowed for 2.7 years, on the average (the age groups were not shown in greater detail). The widowed males in the 20-24 age group had probably been widowed a shorter time, on the average. However, to be conservative, the 2.7-year figure will be used. We therefore estimate that not more than 2.7 times 1,076, or 2,905, males had become widowed at the same ages and times as the 20- to 24-year-old widowed in the Census, and had remarried before the Census. Since this is only 32 per cent of the 9,060 that remained widowed, we estimate that the mortality rate among those males who remained widowed was not more than 32 per cent greater than among all those who had become widowed. It should be emphasized that unduly conservative assumptions have been made regarding duration of widowhood and regarding the selective good health of those widowed who remarried, and therefore the actual effect of the selectivity was probably considerably less than the estimated effect. Similarly, there were 25,000 remarriages in the 7½-year period at ages 14-24 among females previously widowed, from which we estimate that the mortality rate among those 20- to-24-year-old females who had remained widowed was not more than 40 per cent greater than among all those who had become widowed, and was probably considerably less than this. Thus, the estimated maximum effect of the selectivity is slight, relative to the observed tenfold excess risks among those who had remained widowed, compared to the married.

It is difficult to estimate the combined effect of the three artificial factors which have already been discussed. Perhaps in combination they account for a major part of the excess risk observed in the widowed for several causes of death with relatively low excess risk, such as malignant neoplasms and chronic

rheumatic heart disease. If the approximately twofold excess risk of death from these two causes observed in the widowed is assumed to be a measure of the combined effect of these artificial factors, then relative risks of more than fivefold would remain unexplained by these artificial factors for several causes of death previously mentioned.

Several additional hypotheses might be considered as explanations of the observed relationship. Two of these hypothesize that the conditions of widowhood do not themselves produce an increased risk of mortality, but that the event of becoming widowed identifies a group which has an excessive risk for other reasons. The first of these hypothesizes that individuals with a short survival potential tend knowingly or unknowingly to choose mates with a similar potential. This might be called the "mutual choice of poor-risk mates" hypothesis. It is our impression that individuals with certain obvious disabilities of the limbs or sense organs do tend to marry individuals with comparable disabilities, for reasons of compatability. If psychological characteristics are important in the etiology of cerebral vascular disease and arteriosclerotic heart disease, it is not implausible that individuals with those psychological traits likely to produce early death from these diseases would tend to choose mates with similar traits. Such a hypothesis was considered by Ciocco who analyzed the causes of death among a sample of 2,571 husbands and wives who died in Washington County, Maryland, during 1898 to 1938 and found a tendency for husbands and wives to die from the same cause when either tuberculosis, influenza and pneumonia, heart disease, or cancer was the cause of death.[6] This tendency was most striking for the first two causes of death and is obviously explained on the basis of the spread of infection in the family environment. However, for heart disease and cancer, he felt that conscious or subconscious marital selection must be considered as an explanation.

The second of these hypotheses is that the widowed and their deceased spouses shared common unfavorable environmental factors which led to the death of the first spouse and result in an excessive risk to the surviving spouse. This might be called the "joint unfavorable environment" hypothesis.

Finally, there is the hypothesis that the event and new conditions of widowhood themselves have a deleterious effect which results in an excessive death rate. This hypothesis would include the effects of the grief, the new worries and responsibilities, alterations in the diet, work regime, and recreational life, the frequently reduced economic condition, and other specific changes in characteristics resulting from becoming widowed. This will be called the "effects of widowhood" hypothesis.

Although the widowed may tend to be of a lower economic status than married people of the same age, there is some indication that the marked difference in mortality between the young widowed group and the young married group for the diseases of special interest here is not primarily due to an economic difference per se. Lilienfeld[7] has analyzed the variation in mortality from heart disease between socioeconomic classes in Baltimore, Maryland, and in England and Wales. In Baltimore, the census tracts of the city were ranked according to the median monthly rental given in the 1950 Census, and then assembled into

five groups containing approximately equal proportions of the population. The 15,000 deaths due to heart disease to Baltimore residents in 1949-1951 were used to calculate age-sex-race specific death rates, by socioeconomic class of census tract of residence. For arteriosclerotic heart disease (International List No. 420), there was no significant relationship between socioeconomic class and the age-adjusted death rate in either sex or race. For "other myocardial degeneration" (No. 422), the lowest socioeconomic class had an age-adjusted death rate 2.4 times the rate in the highest socioeconomic class for white males and 1.8 times for white females. For hypertensive heart disease (Nos. 440-443), the comparable ratios were 1.3 and 1.8, respectively. For England and Wales, the distribution of the population in the 1951 census by social class and of the 1950 deaths at age 20 to 64 by social class were used to calculate standardized mortality ratios (s.m.r.) by social class. For "coronary heart disease" the s.m.r. for males in the professional group (the highest class) was 1.7 times the s.m.r. for male unskilled workers (the lowest class), while for married females classified according to husband's social class the ratio was 1.2 in the reverse direction. For "myocardial degeneration," the s.m.r. for male unskilled workers was 2.0 times the s.m.r. for male professionals, while for married females the ratio was also 2.0 in the same direction. In all of these data for Baltimore and for England and Wales, the rates for the middle socioeconomic classes tended to be between the rates for the highest and lowest classes.

While the results described above show some inconsistency between males and females, and between Baltimore and England, they have one feature in common which is pertinent to this discussion. The relative risks for those in the lowest socioeconomic class compared to those in the highest class are much smaller than for the young widowed group compared to the young married group in this study. Even if all of the young widowed were in the lowest socioeconomic class and all of the young married were in the highest class, the economic difference per se would not account for more than a minor portion of the observed difference in risk between these two marital status groups, for the types of heart disease considered above.

In actuality, the economic difference between the young widowed and the young married groups is not nearly as great as that assumed above. The issue of *Current Population Reports* previously referred to[5] gives the distribution of urban and rural nonfarm persons of 14 and over with money income, according to marital status in April, 1948, and income in 1947. For total widowed males, 43.9 per cent had an income under $1,000 and 7.4 per cent had an income of $4,000 and over, compared to 8.9 per cent and 18.7 per cent, respectively, for males married once with wife still present. The data are not available on an age-specific or age-adjusted basis. We believe the difference would be much less for comparable age groups, since the age distribution of the widowed would tend to be quite different than for those married once with wife still present, and age is also related to income. These data and considerations further reduce the probable contribution of an economic difference per se to the observed differences in mortality from heart disease between the young widowed and the young married groups.

SUGGESTED STUDIES

Clearly, the relationships and hypotheses discussed should be investigated by more definitive types of studies. Much could be learned by relating deaths among the young widowed to the earlier deaths in their spouses. This might be done on a retrospective basis, similar to but an expanded version of the study by Ciocco,[6] if the informant or attending physician mentioned on the death certificate for the widowed individual could supply the name, place of death, and date of death of the earlier deceased spouse. Then, the death certificate for the earlier deceased spouse could be located. The comparison of the causes of death, the interval of time between the two deaths, the duration of the first deceased spouse's illness, the age difference between the spouses, and the socio-economic status of the residence of each deceased would all be pertinent items obtainable from data appearing on the two death certificates. If the informant on either death certificate could supply the name of the state or the approximate date of marriage of the deceased spouses or both, this might be used to locate the marriage certificate. This would provide the age of each at the time of marriage and the duration of marriage, which might also be pertinent. Due to the relatively small number of deaths among the young widowed, a cooperative study involving several states might be desirable.

If such a retrospective study is not feasible, a prospective study might be an alternative. The informants on the death certificates of young married individuals could be contacted, to obtain a young widowed group. All members of this widowed group could be interviewed at the outset and periodically through a follow-up period. Information could be obtained on physical and emotional characteristics of each spouse at the time of marriage; environmental factors during the marriage, such as diet, exercise, economic condition, work and recreational activities, and sexual and emotional adjustment; emotional response to the deceased spouse's illness and death; and patterns of life during the period of widowhood. The subsequent death rates of this group would be related to these characteristics. The deaths in this group would also be related to the earlier deaths in their spouses. Remarriages would also be identified. Another feature might be a physical examination of the study group as soon after death of the spouse as possible, to determine detectable chronic illnesses present at that time and their onset dates. Physical examinations during the follow-up might also be considered. A control group consisting of a sample of married individuals in the community could be selected, studied, and followed in a similar manner.

Since the death rates in young widowed males are about double the rates in widowed females of the same age, the study might be limited to widowed males under 45. These have an annual death rate of about one per cent, so that a group of 5,000 followed for 4 years should yield approximately 200 deaths. This number would appear adequate to explore the possible hypotheses.

The findings of this mortality analysis are consistent with the results of a morbidity survey carried out in the Eastern Health District in Baltimore, Maryland, during 1938-1943. An analysis of the frequency of various chronic diseases

VOL
10
195

among spouses indicated the presence of a significant association for all circulatory diseases, for hypertensive vascular disease, and for arthritis.[8] Unfortunately, the numbers in the sample were small, which necessitated limiting the analysis to those over 45 years of age. Also, the information was obtained by interview and may be subject to a certain degree of error.

In conclusion, we believe that further study of the marked relative excess in mortality from several leading causes of death among the young widowed group, compared to the young married group, may lead to important findings regarding the etiology of these diseases.

## SUMMARY

The relationship between marital status and mortality appearing in statistical data published by the National Office of Vital Statistics on deaths and death rates in 1949-1951 by marital status, age, color, and sex was reviewed. The main feature was the lower death rate in the married group than in the single, widowed, or divorced, at every age. The outstanding excess risk of mortality among the young widowed was noted. This excess risk in the widowed under age 35, compared to the married, was greater than tenfold for at least one of the specific age-sex groups involved for several leading causes of death, including arteriosclerotic heart disease and vascular lesions of the central nervous system.

Possible explanatory hypotheses were discussed. It seems unlikely that the several possible artificial factors mentioned, including the "selection" factor, could have accounted for more than a small part of the great excess risk noted in the young widowed for these several causes of death. Three biologic hypotheses were suggested in explanation of what appears to be a genuine association between young widowhood and subsequent mortality. These were called the (1) "mutual selection of poor-risk mates," (2) "joint unfavorable environment," and (3) "effects of widowhood" hypotheses.

Further studies, of both the retrospective and prospective types, were suggested to assess these hypotheses.

We wish to express our appreciation to Lise F. Davignon, M.D., M.P.H., for her assistance in the preparation of this paper.

## REFERENCES

1. Mortality From Selected Causes by Marital Status—United States, 1949-1951, Vital Statistics—Special Reports, Vol. 39, No. 7, National Office of Vital Statistics, U. S. Public Health Service, May 8, 1956.
2. Mortality Lowest in Married Population, Statistical Bulletin of the Metropolitan Life Insurance Company **38**:4, 1957.
3. 1950 Census of Population, Vol. II, Characteristics of the Population, Part 1, U. S. Summary, Table 102.
4. Socioeconomic Characteristics of Persons Who Married Between January, 1947, and June, 1954: United States Vital Statistics—Special Reports, Vol. 45, No. 12, National Office of Vital Statistics, U.S. Public Health Service, Sept. 9, 1957, Table 21.
5. Marital Status, Number of Times Married, and Duration of Present Marital Status: April, 1948. Current Population Reports, Series P-20, No. 23, Bureau of the Census, U. S. Department of Commerce, March 4, 1949.
6. Ciocco, A.: On Mortality in Husbands and Wives, Human Biology **12**:508, 1940.
7. Lilienfeld, A. M.: Variation in Mortality From Heart Disease—Race, Sex, and Socioeconomic Status, Pub. Health Rep. **71**:545, 1956.
8. Downes, J.: Chronic Disease Among Spouses, Milbank Memorial Fund Quart. **25**:334, 1947

# THE HISTORICAL DEVELOPMENT OF CLINICAL THERAPEUTIC TRIALS

J. P. Bull, M.D.,
Birmingham, England

*From the Medical Research Council, Industrial Injuries and Burns Research Unit, Birmingham Accident Hospital*

(Received for publication June 15, 1959)

VOL
10
195

DELIBERATE experiments designed to assess upon patients the value of therapeutic procedures are an essential feature of modern medicine. These clinical trials link laboratory results with medical practice and so provide a reliable basis for advance in treatment. They are seldom easy, are usually costly in time and effort, and sometimes are dangerous to patients. The danger may be to patients treated in the trial or to patients badly treated as a result of reliance upon an inadequate trial. The efficiency of these experiments is thus a matter of some importance.

A study of the problems tackled in the past, by what methods, and with what success, should yield lessons for the future. It may show some of the pitfalls and limitations of the clinical therapeutic experiment as well as its triumph in certain circumstances. The conditions for such success may be seen and, further, the best methods to use when those conditions are present.

## THE ANCIENT WORLD

The methods of therapeutic investigation used by the ancient Egyptians are unknown. The outstanding lessons from records discovered are (1) the extreme antiquity of systematic medicine (a papyrus in the British Museum gives prescriptions believed to date from circa 2000 B.C.,[1] the time of the building of the great Pyramids), and (2) the high quality of treatment for certain lesions. While medical conditions were typically treated with a combination of ritual exhortations and fantastic mixtures of herbs and natural products,[2] surgical treatment and, in particular, minor surgery was at a much higher level. The Edwin Smith papyrus,[3,4] dating from circa 1600 B.C., refers to bandaging and stitching of wounds and, for instance, the treatment recommended for a dislocated jaw by reduction is that of modern surgery. We do not know how such methods of treatment were worked out, but it is interesting to note that the

Based on part of an M.D. Thesis: A Study of the History and Principles of Clinical Therapeutic Trials, Cambridge, 1951.

types of lesion so excellently treated were all of simple etiology, mostly accidental injuries, with mechanisms of production and treatment not dissimilar to those of structural engineering of which the ancient Egyptians were such masters. The success of the treatments would be unequivocal, so that therapy based on simple mechanical principles could readily be tested by trial and error. The simplicity of such surgical conditions can be contrasted with the complexity of etiology, diagnosis, and criteria of cure of medical conditions likely to have been prevalent in the Nile valley. Even with modern aids the differentiation of tropical fevers is not easy and one of the characteristic Egyptian diseases, schistosomiasis (possibly the A'a disease of the papyri) has only recently been the subject of satisfactory therapeutic trials.

The surviving records of Babylon and Assyria show a state of medical knowledge similar to that of ancient Egypt[5] and there is much evidence of interchange with Egyptian medicine. An elaborate pharmacopeia was used but much of the therapy was irrational or ritualistic. Exorcism of hostile powers was the declared aim of treatment but incantations were often combined with practical measures such as the use of poultices, purgatives, and enemas. No elaborate trial would be necessary to test the effects of these simple treatments and the persistence of the associated irrational ritual suggests a low level of scientific criticism. Not all the ritual, perhaps, should be so dismissed; it is known, for instance, from excavations that public hygiene was highly developed in Babylon and the frequent recommendation of purification by water or fire may indicate that lessons learned in hygiene were applied in therapeutics.

According to Herodotus[6] it was a custom of the Babylonians to exhibit their sick in a public place so that passers-by might advise suitable treatment based on their experience with similar cases. It seems unlikely that successful treatments would be fully validated on such a system and the records give no suggestion that orthodox medicine benefited by these public clinical trials.

Like that of other ancient civilizations the medicine of the Hebrews was closely allied to religion. Public hygiene, in which the Jews excelled, was administered by the priests and enforced by religious sanctions. The knowledge of preventive measures against infectious diseases was perhaps a product of life in tribal communities, large enough to have an administrative structure and small enough for each member to be known personally. The pressure of hard living conditions and the ever-present danger of communicable disease would favor the growth of a sound knowledge of hygiene. There is no evidence of any deliberate experiments and little is known of ancient Hebrew therapeutics as distinct from preventive medicine. It is possible that the lost book of the *Wisdom of Japheth* which emphasized the value of observation in medicine would have thrown light on methods of advance in treatment.[7]

## GREECE AND ROME

Modern medicine is often said to have begun in Greece. The wealth of original observation and cautious deduction recorded in the Hippocratic books is one of the outstanding achievements of mankind. The study by the Coan

school of the natural history of disease gave rich results in diagnosis and prognosis but there was less success in therapeutics. Surgical treatment, especially of minor complaints, reached a high standard as it had many centuries before among the Egyptians. In contrast, medical therapeutics were dominated by a priori theories. These theories, though less speculative than those of other schools, were oversimple derivations from a general philosophy of nature. Therapy was intended to assist the natural powers of healing by simple exercises and diets. These were rationally deduced from the postulates, but there is no suggestion that experimental trial and judgment by results were used. Thus the *Regimen of Acute Diseases*[8] urges physicians to inquire into the best treatments and goes on to give reasons in favor of barley gruel for fevers, but characteristically the gruel is recommended because it is smooth and soft and not because of its observed effects upon patients, nor is any comparison offered with alternative treatments.

Medicine in contrast to more theoretical subjects has the advantage of enforcing a continual testing of theory by practice. This was recognized by the writer of the *Precepts:* "One must attend in medical practice not primarily to plausible theories, but to experience combined with reason."[9] The natural sequel in therapeutics would seem to be the clinical trial which aims to test the day-to-day impressions and to substitute a deliberate advance for the accumulation of plausible theories. The Hippocratic writers with their confidence in general philosophy did not draw this conclusion. A study of the successors of the Hippocratic school emphasizes different aspects of this paradox between a stated reliance upon observation and an actual trust in a priori theory. The Dogmatists developed the theory of humors and qualities into a rigid formalism.[10] The Empirics, on the other hand, discounted theory and relied upon practical tests only. In this way knowledge of drugs was advanced and poisons and their antidotes, in particular, were widely studied. Attalos and other rulers tested the effect of poisons upon criminals by a kind of reversed clinical trial, and the famous antidote of Mithridates comes from this era. Extreme Empiricism with its undue emphasis on "cures" and its neglect of general principles of etiology and diagnosis led to extravagant polypharmacy and hindered therapeutic progress.

Roman medicine derived directly from that of Greece, and many of its leaders were of Greek origin. The arguments of the conflicting schools continued but the outstanding men favored an eclectic compromise. Celsus, after reviewing the history of classical medicine, states fairly the Empiric view of the development of therapeutics, ". . . careful men noted what generally answered the better, and then began to prescribe the same for their patients. Thus sprang up the Art of Medicine which from the frequent recovery of some and the death of others, distinguished between the pernicious and the salutary."[11] Celsus agrees that ". . . nothing adds more to a really rational treatment than experience" and concludes that ". . . the Art of Medicine ought to be rational but to draw instruction from evident causes, all obscure ones being rejected from the practice of the Art, although not from the practitioner's study." Like other early writings *De Medicina* cites little evidence to support the claims made for its treatments

and though these include sound use of some drugs, in particular local application of astringents, many others are only "sympathetic" remedies and it seems unlikely that either the experience of "careful men" or consideration of "evident causes" could justify giving ox spleen for splenic enlargement or pole reed for injuries by splinters.

In the first century A.D. Dioscorides wrote the earliest scientific account of medical botany.[12] He recognized natural families of plants and classified their medicinal qualities. His work was the foundation of herbals for 16 centuries,[13] but he did not describe methods of testing therapeutic action. He recommended Mandragora wine for insomnia and as an analgesic draught which could be used in surgical operations. Dioscorides was himself a military surgeon and had, presumably, opportunities for trying such analgesic properties. Some other important pharmacologically active plants, for example, digitalis and Atropa, are not mentioned and others, such as Salix, are recommended for a multitude of complaints often not including what is now known to be their most valuable action. It would be unreasonable to expect an exhaustive account in a pioneer work but the absence of these important parts of materia medica may explain some of the impotence of medicine in medieval times when Dioscorides was followed slavishly.

The advances in surgical technique in Roman times were probably related to experience with gladiatorial and military injuries.[13] Galen himself served as physician to the gladiators and became one of the great founders of scientific medicine. His experimental approach was well shown in physiology but, although he tried to deduce therapeutics rationally from knowledge of disease and understanding of remedies, inadequate testing permitted general theories such as that of treatment by contraries to override practical results upon patients.

### THE MIDDLE AGES

After the fall of the Roman Empire the scientific trend in European medicine was arrested. The newly developing Moslem culture took up and preserved the classical and early Christian medical teaching. The doctrines of Islam favored conservatism in therapeutics and, in particular, the objection to touching the human body delayed advances in anatomy and surgery.[13] The authority of Galen was, in general, accepted and his teachings were often mixed with astrology. Arab merchants dominated the spice and drug trade so it is not surprising that there were numerous pharmacologic experiments.[14] Rhazes (860-932) was outstanding in this field and Avicenna (980-1037) in his encyclopedic *Canon* gives some interesting rules for the testing of drugs. He suggests that in the trial of a remedy it should be used in its natural state upon uncomplicated disease, that two opposed cases be observed, and that study be made of the time of action and of the reproducibility of the effects. He further stated: "The experimentation must be done with the human body, for testing a drug on a lion or a horse might not prove anything about its effect on man."[15] These rules imply a very modern approach but there seems to be no record of their detailed application. It is possible, however, that documents not yet translated may throw more light on this.

In Europe the supernaturalism of medicine of the monastic period was countered in the eleventh and twelfth centuries by the School of Salerno. Here a simple rational therapy was based on direct study of disease. The famous *Regimen sanitatis* gives sensible advice on hygiene and diet and considerable detail on blood letting but no suggestion of methods of testing remedies.[16] Roger of Palermo in his *Practica* (circa 1170) recommended seaweed for goiter and mercury salves for skin diseases, and one can only surmise that some kind of simple trial had shown their usefulness.[13]

Latin translations of Arabic scholars such as Avicenna became available in the thirteenth century and, at the same time, the Roman Church which dominated intellectual life was rejuvenated by the Franciscan movement. An outstanding exponent of this medieval renaissance was the Franciscan Roger Bacon (1214-1292). In his *De erroribus medicorum*[17] he points out inconsistencies in current medical teaching but excuses some of the defects "for it is exceedingly difficult and dangerous to perform operations on the human body, wherefore it is more difficult to work in that science than in any other. So that physicians are always to be excused since needs must be that they have deficiencies. For the operative and practical sciences which do their work on insensate bodies can multiply their experiments till they get rid of deficiency and errors, but a physician cannot do this because of the nobility of the material in which he works; for that body demands that no error be made in operating upon it, and so experience (the experimental method) is difficult in medicine. Wherefore physicians are to be excused for their defects more than are workers in the sciences." Bacon favored appeal to experience and mathematical demonstration but does not appear himself to have applied these principles to therapeutics since his own recipes given at the end of *De erroribus* are on traditional lines.

In the fourteenth and fifteenth centuries the attempt was made rather to reconcile Aristotelian dialectics with Arabian medicine than to continue the lead of the thirteenth century theorists by advancing knowledge by experiment. This failure was in keeping with the medieval intellectual atmosphere and its emphasis on faith, authority, and philosophical idealism,[18] but quite apart from the intellectual controversies the practitioners of many crafts continued to make sound observations and steady progress. The use of water power was greatly developed, methods of agriculture improved, and, with the introduction of the modern type of ship's rudder, long sea journeys became possible.[19] Similar progress presumably occurred in therapeutics. Something of this is suggested by Henri de Mondeville (1260-1320) who defended cleanliness in the management of wounds against the "coction" and "laudable pus" of the Arabian commentators on Galen.[20] His claim that "wounds dry much better before suppuration than after it" implies direct observation and appeal to experience rather than to authority.

## THE RENAISSANCE

It is not appropriate to discuss here the various aspects of the revival of learning which, beginning in Italy in the fifteenth and sixteenth centuries laid the foundation of modern science. The renewed study of Latin texts by Petrarch

(1300-1374) and his successors, of Greek originals by Boccaccio (1313-1375) and others, the invention of gunpowder (circa 1330) and of printing (1440-1450), the fall of Constantinople (1453), the discovery of America (1492), and Magellan's circumnavigation of the world (1519-1522) all played a part in this enormous development of theory and practice. Two contrasting threads of this story are important in the history of clinical trials. In common with all other branches of thought, medicine was affected by the revival of Classical Humanism. Early medical Humanists such as Leonicenus (1428-1524) and Linacre (1460-1524) provided new and accurate translations of Hippocrates and Galen; they and their followers were thus in an excellent position to attack contemporary medical teaching based on indirect annotations of the classics, but their criticism of the new empirical medicine could be just as stringent. This empiricism was a continuation of the mixed science and magic of the alchemists. Paracelsus (1493-1541), its great exponent, was a pupil of Leonicenus and from him acquired a lifelong respect for Hippocrates, but his public burning (1528) of the works of Galen and Avicenna shows his violent opposition to current medical orthodoxy. He wrote and spoke in the language of the common people and from them collected information on folk medicine. Being also an expert alchemist he made experiments in the use of metals in therapy. No details are available of his methods of trial, but he says, "experience has shown that Mercury is the sovereign and only remedy for the cure of all ulcers tainted with the great pox,"[21] and again, "practice should not be based on speculative theory; theory should be derived from practice. Experience is the judge; if a thing stands the test of experience it should be accepted; if it does not stand the test it should be rejected."[22] The implied scientific approach was not always sustained. Paracelsus also originated one of the earliest sympathetic remedies. The response of the Humanists to these new ideas was decidedly hostile. One of them, Fernel (1497-1558), also an admirer of Hippocrates, attacked the use of mercury not because it was ineffective but because of its empirical origin. Fernel based his own therapeutics on the doctrine of contraries and favored guaiac for syphilis.[23]

Meanwhile Leonardo da Vinci (1451-1519) had developed the theory and practice of the modern scientific experiment. He saw the importance of mathematical demonstration and said, "I shall test by experiment before I proceed further, because my intention is to consult experience first and then with reasoning show why such experience is bound to operate in such a way. And this is the true rule by which those who analyse the effects of nature must proceed; and although Nature begins with the cause and ends with the experience, we must follow the opposite course, namely . . . begin with the experience and by means of it investigate the cause."[24] Although there are claims to the contrary, it seems doubtful that Leonardo appreciated the full scope of scientific method. His experimental approach continued to be that of the craftsman and engineer—the testing of materials and of ad hoc hypotheses rather than part of an inductive system.[25] Although the testing of drugs by clinical trials falls well within the scope of his type of experimentation Leonardo did not apply these methods to therapeutics and was scornful of doctors, probably with justification. The few recipes he gives in the *Notebooks* are of traditional type.[26] Ambroïse Paré (1510-

1590) the greatest figure in Renaissance surgery, popularized the revolutionary anatomic teaching of Vesalius (1514-1564) and also made one of the earliest reported clinical trials, albeit an unintentional one. In 1537 while serving with the Mareschal de Montegni he was responsible for the treatment of the wounded after the capture of the castle of Villaine. They were so numerous that, he says, "at length my oil lacked and I was constrained to apply in its place a digestive made of yolks of eggs, oil of roses, and turpentine. That night I could not sleep at my ease, fearing that by lack of cauterization I would find the wounded upon which I had not used the said oil dead from the poison. I raised myself early to visit them, when beyond my hope I found those to whom I had applied the digestive medicament feeling but little pain, their wounds neither swollen nor inflamed, and having slept through the night. The others to whom I had applied the boiling oil were feverish with much pain and swelling about their wounds. Then I determined never again to burn thus so cruelly the poor wounded by arquebuses."[27] The other recorded instance of Paré's use of such methods of trial is perhaps less fortunate. One of his later masters, Charles IX had a bezoar stone supposed to be a universal antidote. Paré criticized this claim and suggested it be tried on a convict. A prisoner agreed and instead of being hanged was given poison and the bezoar stone. He died after about 7 hours and Paré did an autopsy which confirmed that death was due to corrosive sublimate. Similar tests of antidotes upon criminals were made by other rulers about this time, and in view of the commonness of poisoning such ruthlessness was perhaps understandable.

During the sixteenth century there was an increasing interest in natural history, stimulated by the numerous voyages of discovery. There was a corresponding development of systematic botany and the work of Fuchs (1501-1566) and Valerius Cordus (1515-1544) provided first-hand descriptions and classification of plants.[28] This resulted in improvements in materia medica and the earliest pharmacopeias date from this time. There were also several treatises on purges and presumably some kind of testing of pharmacologic properties was performed, aimed at symptomatic rather than curative therapy.

## THE SEVENTEENTH CENTURY

The seventeenth century saw great growth of the application of scientific method, but medicine was less affected by this than were physics, chemistry, and biology, and there was little development of therapeutic trials. As in the preceding periods practical men did note and learn from comparative observations. For example in the first expedition to India by the newly formed East India Company in 1600, there were four ships; on one of them only, that of General James Lancaster, was lemon juice provided and this ship was almost free from scurvy whereas the others were badly affected during the slow voyage. The Company was sufficiently far-sighted to supply all its ships with lemon juice for subsequent voyages.[29] This lead was not taken up by medical men for other diseases; in fact, it was not until the next century that a corroborative trial of this same treatment for scurvy was made.

Francis Bacon (1561-1626) the great protagonist of the inductive method in science, devoted a section of his *De augmentis scientiarum* to medicine. He discussed the dangers of crude empiricism to which doctors were driven by the demand for "cures" and said of therapeutics, ". . . this part of phisic which treats of authentic and positive remedies, we note as deficient; but the business of supplying it is to be undertaken with great judgment, and as by a committee of physicians chose for that purpose."[30] How this "committee" was to proceed Bacon did not specify.

The outstanding contribution to physiology made by Harvey's *De motu cordis* (1628)[31] inspired others to attempt overfacile simplifications of medicine based on general principles analogous to those discovered in physics by Newton. A similar result was produced by Boyle's brilliant chemical discoveries which led in the later iatrochemical school to speculations which were often far removed from sound scientific principles.[32] There was failure to apply the applauded scientific methods to test hypotheses; instead reliance was still placed in a priori reasoning. Boyle (1627-1691) was himself very skeptical of current medicine though very interested in drugs and diseases. He made numerous suggestions for imitating medicinal waters and cheapening drugs and in *The Usefulness of Natural Philosophy* he says, "Another way by which the naturalist . . . may help to lessen the changeableness of cures is by showing there hath not yet been sufficient proof of their having any medicinal virtues at all. . . ."[33] Boyle also recommended experiments on animals and with Wren he performed the early blood transfusions. In 1657 they succeeded in transfusing animal blood into a human subject as is recorded by Pepys, but this was hardly intended as therapy and when extended as such by Denis in France it was soon discredited.[34]

An outstanding new therapeutic agent introduced to Europe about this time was Cinchona bark. The precise circumstances of its discovery are still disputed but it is clear that it had been used for many years previously by the Peruvian Indians before it came to the attention of Spanish colonizers about 1630. Its success in the treatment of fevers must have impressed the Jesuit missionaries for they brought the bark to Europe in 1632. It had a hostile reception by orthodox medicine. This was partly due to suspicion of the Jesuits but a further objection was that its acceptance meant an overthrow of Galenical theory. As Ramazzini later observed, cinchona did for medicine what gunpowder had done for war.[13] No careful trials of the value of this drug seem to have been made and there is even doubt as to the nature of some of the bark used since, apart from fraudulent preparations, there seems to have been confusion between cinchona with its antimalarial action and quinaquina, the source of Peruvian balsam.[35] Jesuits' bark, discounted by doctors, became the property of quacks. One of these, Talbor (1641-1681), achieved great fame with his remedy for fevers and became Physician to Charles II.[36] Sydenham (1624-1689) broke away from orthodoxy and, although a nihilist in many fields of therapy, used and recommended Peruvian bark for intermittent fevers. His methods of trial are unfortunately not known, but his skepticism of the newly developing sciences of anatomy and physiology[37] make it unlikely that his techniques were derived from the contemporary scientific trends of the Royal Society. His reliance upon

clinical experience contrasts with Boyle's confidence in experiment. The marriage of these two approaches to produce a satisfactory clinical trial did not occur in the seventeenth century, although it appears that Locke (1632-1704), a friend of both Boyle and Sydenham, had something of this sort in mind in his projected work *Ars medica* on the philosophy of medicine.[38] Therapeutics, instead, was at the mercy of fashion and quackery. The status of antimony provides an example; it had first been publicized at the beginning of the century by extravagant claims made in *The Triumphal Chariot of Antimony*. This book was probably written by Thold but was attributed to a fifteenth century monk, "Basil Valentine." By the middle of the century antimony had fallen into disrepute, but in 1657 it was used in treating an illness of Louis XIV. The king recovered and antimony again became popular for the treatment of fevers.[39]

By contrast, surgical treatment made more steady progress. Wiseman (1622-1676), the outstanding English surgeon of the century, accepted as self-evident the virtues of bleeding, purging, vomiting, sweating, and salivation but nevertheless he made acute observations on simple therapeutic advances. He vividly describes successful treatment of edema of the legs with laced stockings, the recurrence of the trouble when treatment was interrupted, and the further improvement on re-placing the appliance.[40] This kind of simple trial with limited mechanical aim has frequently contributed to practical surgical techniques.

## THE EIGHTEENTH CENTURY

In the eighteenth century the new advances in scientific knowledge and method began to be applied to therapeutics. This was achieved often in the face of opposition from the orthodox theorists who still tried to impose a priori systems upon medical treatment. It is not surprising therefore to find that the advances were made sporadically by persons of independent mind often working in the provinces away from the fashionable systematists. In default of a scientific background new treatments were borrowed from folk medicine but the need of adequate trials of their effects was increasingly realized.

Inoculation as a preventative of smallpox was introduced from Constantinople by Maitland (1668-1748) and Lady Mary Wortley-Montague. They persuaded King George I to permit a trial upon six Newgate convicts in 1721.[41] All survived the operation and were released; one, in whom the inoculation failed, was found to have had smallpox before; another was exposed to infection after the treatment and was found to be immune. The results were thought to be conclusively in favor of inoculation and it became widely practiced. The trial has since been criticized because it was uncertain which of the subjects had previously suffered from the disease and the trial would seem to have tested the safety of the procedure rather than its effectiveness.[42] Subsequent experience soon showed that inoculation could have a considerable mortality and that complete protection could not be assured. These variable results were obtained with different modifications of technique which the early trial could not have been expected to test, but the numbers treated and the precision of the trial were inadequate to give a fair picture of the effects of the operation.

Other treatments continued to be proposed on very slender evidence. Dover published his *Ancient Physician's Legacy to His Country* in 1733,[43] claiming extravagant cures of gout, dropsy, and diabetes, and in the same year Bradley published a rejoinder accusing Dover of quackery[44]; neither performed careful trials of the recommended treatments. Cures for the stone were in great demand and a Mrs. Stephens was given a Parliamentary grant of £5,000 for her secret remedy. The very able experimental physiologist Stephen Hales (1677-1761) was on the committee which examined the claims for the treatment. Patients supposedly cured were questioned and the grant was supported. Hales may have had some misgivings since he later experimented with the mixture and decided that only the "soap lees and lime of eggshells" which it contained among many other ingredients could have had any solvent effect upon calculi[45]; he does not seem to have considered the necessity of careful trials upon patients. Heberden (1710-1801) was critical of many traditional treatments and in his *Antithriaka* (1745) he dealt the final blow to the spurious antidote of Mithridates famous in various forms for 2,000 years.[46] He based his attack on the variable constitution and the inconsistencies and unreasonableness of the supposed properties of theriac rather than upon a demonstration of its ineffectiveness. Another popular remedy was publicized in 1744 by Bishop Berkeley in *Siris: A Chain of Philosophical Reflexions and Inquiries Concerning the Virtues of Tar Water.*[47] From anecdotal evidence of tar water curing smallpox and scurvy he goes on to quote his own use of it in the Irish epidemics; the transition is then made through chemistry and physics to philosophy and theology. Although no comparative results are given to substantiate the claims, tar water became so popular that a special warehouse was opened in London for its distribution.[48] Hales also examined this remedy and cautiously attributed discrepancies in its effects to variations in method of preparation.[49]

At this time an investigation of quite a different quality was made by a ship's surgeon, James Lind (1716-1794). Appalled by the ravages of scurvy in Anson's recent circumnavigation of the world when three-quarters of the men died from the disease, Lind planned a comparative trial of the most promising scurvy "cures." On May 20, 1747, he says, "I took twelve patients in the scurvy, on board the Salisbury at sea. The cases were as similar as I could have them . . . they lay together in one place . . . and had one diet common to them all." To two of them was given a quart of cider a day, to two an elixir of vitriol, to two vinegar, to two oranges and lemons, and to the remaining two "an electuary recommended by an hospital surgeon." "The most sudden and visible good effects were perceived from the use of the oranges and lemons, one of those who had taken them being at the end of six days fit for duty. . . . The other . . . was appointed nurse to the rest of the sick."[50] Apart from the cider, which seemed to do a little good, the other remedies were ineffective. In spite of this apparently conclusive demonstration Lind himself continued to cling to other theories; in discussing the treatment of scurvy later in the book he recommends fruit and vegetables but he gives priority to "pure dry air" and says again "hence the first step . . . is change of air." Others seemed no less doubtful and there were several suggestions for cheap alternatives supposedly equal to fruit. A

VOL.
10
1959

conclusive trial of fruit as an antiscorbutic is sometimes attributed to Captain Cook who used it on his second voyage in 1772; but this cannot be sustained since many different prophylactics were used, among them beerwort, of which Cook wrote to Sir John Pringle that it was "one of the best anti-scorbutic sea-medicines yet found out."[51] The British Navy did not supply lemon juice to its ships until 1795.

A close association existed at this time between general science, exploration, and medical research. Pringle was one of the founders of preventive medicine and, when awarding Cook the Royal Society's Medal, praised him as much for his achievement in preventing scurvy as for his geographic discoveries. Sir Joseph Banks, who succeeded Pringle as President of the Royal Society, had himself accompanied Cook on his first voyage around the world and became a great supporter of scientific work; among his many correspondents were John Hunter and Edward Jenner. Another feature of the intellectual life of the eighteenth century was the rise of medical and scientific societies where the members met to discuss a wide range of subjects. The Lunar Society in Birmingham included the leaders in science, engineering, and medicine of that rapidly growing industrial area. Contact was maintained with continental and American workers, and in this atmosphere of scientific triumphs it is not surprising that along with genuine advances less authentic ones were introduced. Many experiments were made on static electricity, and electric shock treatment became popular under the name of "Franklinism." How little this name was justified is shown by Franklin's letter to Pringle in 1757 describing his tests of electricity upon paralyzed patients who came to request it. The letter recounts the temporary improvement which Franklin cautiously suggests may have been due to the exercise of the journey in coming for treatment or the "spirits given by the hope of success."[52]

Similar evidence of popular and even medical credulity was shown by the success of Perkin's tractors. These were metallic rods which were supposed by some electrical influence to cure a great variety of diseases. The treatment was recommended by several distinguished doctors, including Nathan Smith, the founder of Yale Medical School, and an "Institute of Perkinism" was founded in London.[39] Haygarth (1740-1827) of Bath submitted the method to a devastating clinical trial described in his book *Of the Imagination as a Cause and as a Cure of Disorders of the Body*.[53] On 5 patients he used imitation "tractors" made of wood and all but one of the patients were relieved. The following day he repeated the treatment using instead a pair of genuine tractors and obtained identical results. Haygarth aptly quotes Lind's comment on the fictitious scurvy remedies used at the siege of Breda, "an important lesson in physic is here to be learnt, viz., the wonderful and powerful influence of the passions of the mind upon the state and disorders of the body. This is too often overlooked in the cure of diseases. . . ."

Several new medical schools were founded in the eighteenth century, and another important development was in hospital construction. This, after having "approached perfection in the fifteenth century,"[13] had been seriously neglected in the sixteenth and seventeenth. The number of hospitals actually diminished

as a result of the disbanding of religious houses during the Reformation, and little was done to replace the loss in spite of increasing populations. In the eighteenth century the position began to be corrected by the founding of new hospitals both in London and in the growing provinicial cities. William Withering (1741-1799) became Physician to the Birmingham General Hospital at its foundation in 1778. Here were more patients upon which to continue his trials of the preparations of digitalis first brought to his notice as a constituent of a folk remedy for dropsy. In his *Account of the Foxglove*[54] he gives his experience of the drug in 163 cases. There is a masterly description of his establishment of correct dosage. At first, he thought it necessary "to bring on and continue vomiting," then he persisted until nausea was caused, later he aimed at either diuresis, sickness, or purging, but having noticed slowing of the pulse in some cases his final method was not to repeat the drug too quickly and to continue "until it either acts on the kidneys, the stomach, the pulse, or the bowels." Withering's achievements in deciding the types of patients who would benefit from digitalis are equally remarkable when it is remembered that virtually nothing was then known of the pathology of different kinds of edema. After describing patients apparently with hypertension and malignant ascites as showing little improvement he says, "on the contrary if the pulse be feeble or intermitting, the countenance pale, the lips livid, the skin cold, the swollen belly soft and fluctuating, or the anasarcous limbs readily pitting under the pressure of the finger, we may expect the diuretic effects to follow in a kindly manner." In assessing the value of this treatment, Withering relied upon close observation of his patients and used all the appropriate clinical methods then available; he compared the results with the same patients' previous condition and sometimes with their relapses on discontinuing the drug. This technique is particularly appropriate for such a problem.

The originality of Withering's discovery was disputed by his fellow member of the Lunar Society, Erasmus Darwin,[55] but Darwin's sporadic use of the drug in apparently excessive dosage bears no comparison with the other's systematic study of a long series of patients.[56] Another acquaintance of Withering's was Thomas Fowler who had succeeded him at Stafford Hospital (founded 1772). Here Fowler made his study of arsenic solutions in the treatment of a variety of diseases. In his *Medical Reports of the Effects of Arsenic*,[57] he gives a good account of toxic effects and claims that arsenic is a possible substitute for Peruvian bark in the treatment of ague, and he quotes 247 cases of this disease treated by his method. He claims to have cured two-thirds of these patients, but it appears that his methods of observation were unequal to the difficult task of assessing a remedy for such a notoriously intermittent and relapsing disease, which even with the modern advantages of microscopic diagnosis has only recently been the subject of satisfactory clinical trials.

The history of surgery in this period is dominated by the work of John Hunter (1728-1793), and it was he who established modern surgery on a scientific basis. There is an incident in his early days as a military surgeon during the Belle Isle campaign (1760-1761) which is reminiscent of the experience of Paré two centuries earlier. It was current practice to search for and if possible remove

VOL.
10
1959

the missiles from gunshot wounds. Hunter describes 5 cases in which this was not done and explains that the "neglect rather arose from accident than design." The patients were Frenchmen who had hidden since being wounded and, with only superficial dressings, they all recovered. Hunter modified his treatment of gunshot wounds accordingly and writes, "This practice has arisen from experience, for it was found that balls when obliged to be left, seldom or ever did any harm when at rest and when not in a vital part."[58] As with Paré and the cauterization of wounds, the necessities of military surgery forced this unintentional trial upon Hunter. Being outstanding men, they both drew the correct lessons from these experiences.

At the end of the eighteenth century Jenner published his famous studies on vaccination.[59] This again was a procedure derived from folk medicine and had been suggested by the belief among country people that infection with cowpox prevented subsequent attack by smallpox. Jenner's original account cites 14 persons who, having had cowpox, did not take smallpox when inoculated subsequently. In a further 10 patients, artificial infection with cowpox is described, and 4 of these did not take smallpox when inoculated. A reasonable prima facie case was thus provided in favor of vaccination but the trial was not so conclusive as has sometimes been claimed. The method was the uncertain one of arm-to-arm infection, the people in whom protection was demonstrated were few; this protection was against artificial inoculation and not the naturally occurring disease, and no account was taken of the possibility of natural or previously acquired immunity. George Pearson in 1798 published another smaller trial in which he made a detailed study of the results of inoculating 5 persons with smallpox, of whom 3 previously had had cowpox and 2 had not. The results confirmed the protection given by cowpox but Pearson was guarded in his conclusions and recommended "well-directed observation in a thousand cases of inoculated cowpox."[60] Studies of this magnitude were soon available since smallpox and inoculation hospitals provided ample material,[61] but the conditions of observation and the techniques whereby mixed infections of cowpox and smallpox were transmitted prevented the drawing of clear and consistent conclusions. A smaller trial, the counterpart of Pearson's, was made by Waterhouse, the first doctor to use the method in America.[62] He vaccinated 19 boys; 12 of these he afterward inoculated with smallpox, as he did 2 others who had not been vaccinated. Only these latter 2 took the smallpox infection.

A study of the relative merits of different treatments for syphilis was made by John Pearson. He dedicated his book to Thomas Fowler and in it compared the claims and his experience of various herbal and chemical remedies "to ascertain whether any other substance than Mercury be a true and certain antidote."[63] As surgeon to the London Lock Hospital he had a wide knowledge of the disease and he tried out, apparently not very systematically, any likely remedies. He gives the details of studies on 31 patients in support of his opinion that guaiac, China root, sarsparilla, and other treatments recommended as alternatives were only of value when used in addition to mercury. As to mercury itself, he states that its effectiveness was demonstrated in "not less than twenty thousand cases" of which he had personal experience. He was not blind to its

disadvantages and looked forward to further discoveries since "it were highly desirable to acquire a medicine equally potent as an antivenereal, and not possessing certain active properties peculiar to that mineral." This study, based on simple clinical assessment of results, seems to have been highly successful, particularly so since there was at that time little understanding of the pathology of the disease nor were there accurate tests of cure and, as Pearson was aware, symptoms due to syphilis were frequently confused with those of excessive treatment with mercury.

### THE NINETEENTH CENTURY

Early in the nineteenth century a further emphasis began to be placed on the need for a careful statistical approach in the evaluation of remedies. This was chiefly directed to criticism of extravagant claims. Cobbett, for instance, in his pamphlet *The Rush Light*[64] threw doubt upon Rush's evidence for the value of bleeding and purging in yellow fever, and was one of the first to appeal for the application of statistics to such problems.[32] Theoretical statistics was also advancing rapidly; the *Théorie Analitique de la probabilité* appeared in 1810 and in it Laplace, after reviewing the application of statistical methods in many fields, says, "La même analyse peut-être étendue aux divers résultats de la Médicine."[65] This approach was developed by P. C. A. Louis (1787-1872) chiefly to establish the diagnostic features and the natural history of diseases such as typhoid fever and phthysis.[66] In his *Essay on Clinical Instruction*,[67] he lays down admirable rules for the use of his "Numerical Method" in the assessment of therapy. "As to different methods of treatment, if it is possible for us to assure ourselves of the superiority of one or other among them in any disease whatever, having regard to the different circumstances of age, sex, and temperament, of strength and weakness, it is doubtless to be done by enquiring if under these circumstances a greater number of individuals have been cured by one means than another. Here again it is necessary to count. And it is, in great part at least, because hitherto this method has been not at all, or rarely employed, that the science of therapeutics is still so uncertain; that when the application of the means placed in our hands is useful we do not know the bounds of this utility." Louis goes on to detail some of the necessary precautions, "in order that the calculation may lead to useful or true results it is not sufficient to take account of the modifying powers of the individual; it is also necessary to know with precision at what period of the disease the treatment has been commenced; and especially we ought to know the natural progress of the disease, in all its degrees, when it is abandoned to itself, and whether the subjects have or have not committed errors of regimen; with other particulars." The method was not likely to be easy. "The only reproach which can be made to the Numerical Method . . . is that it offers real difficulties in its execution. For, on the one hand, it neither can, nor ought to be applied to any other than exact observations, and these are not common; and on the other hand, this method requires much more labour and time than the most distinguished members of our profession can dedicate to it. But what signifies this reproach, except that the research of truth requires

much labour, and is beset with difficulty." The most famous example of Louis's use of this method is in his *Recherches sur les effets de la saignée.*"[68] He studied the effects of bleeding upon 78 cases of pneumonia, 33 cases of erysipelas, and 23 cases of inflammation of the throat; he found no appreciable difference in mortality or in duration or type of symptoms or signs between patients bled and those not bled or between patients bled at different stages of the disease. This result, which was contrary to orthodox teaching of the time, caused an uproar among French physicians, but it dealt a fatal blow to the "depletive" treatment then in vogue.

A quite different approach to research in therapeutics was also developed in the early nineteenth century. This was based on tests in animals and its great exponent was Magendie (1783-1855). Similar methods had been used sporadically for many years but now many new chemical substances were being prepared and many alkaloids identified, particularly by French scientists. Magendie and his co-workers tested these substances upon animals to investigate their toxicities and pharmacologic actions. In this way halogen compounds, strychnine, emetine, quinine, morphine, and other drugs were introduced to medicine. Magendie, in his preface to the *Formulaire* states, "la manière d'agir des médicamens et les poisons est la même sur l'homme et sur les animaux"[69] and, with confidence in this belief, trials upon man were limited to confirmatory tests, upon healthy persons, of the pharmacologic actions observed in animals. Such an approach is highly appropriate for drugs used to obtain a physiologic response or for symptomatic treatment but it is inadequate for the assessment of curative value. Pharmacologic investigation is a valuable and perhaps essential prerequisite for a clinical trial but it cannot be a satisfactory substitute.

The rapidly increasing knowledge of chemistry, physiology, and pharmacology made the need for good clinical trials increasingly evident. Several attempts were made to clarify the requirements of a satisfactory trial. An American physician, Elisha Bartlett (1804-1855), in *An Essay on the Philosophy of Medical Science*, which he dedicated to P. C. A. Louis, said that in therapeutic investigations cases which are to be compared must have equal disturbing factors of location, social class, and the like; they should be susceptible of a clear and positive diagnosis; there must be no selection of cases; and the method of treatment must be clearly defined. The certainty of results, he said, will be "in proportion to the fixed and uniform character of the compared facts and to the greatness of their numbers." Bartlett was well aware of the dangers of unbalanced use of the method. Countering the accusation that it tended to therapeutic nihilism, he points out that certain orthodox procedures though they had received no formal trial had been "established by a series of observations of such vast extent as to compensate in a good degree for the absence of the other conditions." It is interesting to note that in this category along with quinine for intermittent fever, opium for colic, and calomel for syphilis he includes bleeding for acute pleurisy, although he warns that "even in these cases it is only by a faithful adherence to the rules and methods which have been described that the exact value of the several remedies can be ascertained." Bartlett dealt also with the danger of losing sight of the individual in the statistical group, "No

acquaintance, however perfect, with the laws of pathology and therapeutics, can ever remove, or in any degree diminish, the necessity of a thorough and discriminating study and knowledge of the single instances which unite to make up the materials of the law."[70]

The influence of Louis was particularly important in the development of scientific medicine in America[71]; one of his admirers, Oliver Wendell Holmes (1809-1894), in a lecture given in 1860, analyzed the errors which led to overmedication. He cited incapacity for sound observation, inability to weigh evidence, the counting of only favorable cases, the "post hoc ergo propter hoc" fallacy, and failure to learn from experience. Holmes further laid part of the blame on the public which "insists on being poisoned."[72]

Clinical trials on the lines of Louis's suggestions were made in Great Britain by Bennett in Edinburgh and Sutton in London. Bennett in 1865 reported his experience, over 20 years, with 129 cases of pneumonia treated on "restorative" principles. He analyzed his patients carefully in respect to sex, age, severity, mortality, and duration of illness and was able to compare his results favorably with those of other series treated by more heroic methods. He also showed that his treatment gave good results in the hands of others and concluded that bleeding and the use of tartar emetic should be abandoned.[73] Also in 1865 Sutton published a trial made upon patients with rheumatic fever under the care of Sir William Gull. At the time, claims were made for a multitude of treatments for this condition and the trial consisted of the careful observation of 20 patients receiving only mint water. The results demonstrated the great natural variation in the disease and the marked tendency to spontaneous curve. Sutton concludes, "A perusal of the above cases tends to show that the best treatment for Rheumatic fever has still to be determined, and will also convince the reader . . . that it is absolutely necessary to understand the natural progress of the disease before any conclusion can be arrived at concerning the operation of remedies. The cases show that too much importance has been attached to the use of medicines, especially in those acute cases where the tendency to a natural cure is the greatest."[74]

Parallel with these developments in medicine, surgery was making enormous strides since the introduction of general anesthetics (1842-1847). The precise priorities of the discoveries are still disputed but it is clear that all the early trials of anesthetics by Long, Wells, Morton, and Simpson were made on very few patients.[75] The unconsciousness of the patient and his subsequent recovery were sufficiently evident not to demand formal control cases nor large series. As it happened, none of the early agents had marked long-term ill effects nor was individual variation of great importance, so the small numbers and short period of study of the early trials did not lead to serious error on these scores. Research into the best methods of administration was made by Snow (1813-1858), who in 1858 published *On Chloroform and Other Anaesthetics*.[76] Here he summarized a wealth of observations both experimental and clinical which could only have been made by someone who, like Snow, combined a scientific approach with the great experience in one field made possible by specialism.

The other great development of nineteenth century surgery was in antiseptic and aseptic methods. Lister (1827-1912) introduced his new techniques initially as rational procedures to avoid sepsis and supported his arguments by the evidence of a few carefully followed cases. In 1870 he published his statistics for amputation operations and compared 35 cases before the use of antiseptics with 40 cases treated by the new method. He showed a mortality of 43 per cent in the former and 15 per cent in the latter but was diffident about drawing conclusions saying, "These numbers are, no doubt, too small for a satisfactory statistical comparison. . . ."[77] Our comment now might be that the numbers are not at fault for the chi square test shows them to be highly significant; what is more open to question is the adequacy of the comparison with previous experience since so many relevant features such as selection of cases for operation must also have changed. Had it been possible, a careful comparative trial of rival methods at this stage might have prevented the bitter and profitless controversy which raged for many years on the subject of the importance and technique of prevention of infection at operation.

The increasing knowledge of bacteriology led to many new applications of immunity to therapeutics. The earliest trial was of Pasteur's vaccine for the prophylaxis of anthrax in animals. The experiment of 1881 is vividly described by the *Times* correspondent[78]: 60 sheep were used; 25 were inoculated and infected, 25 were infected only, and 10 were neither inoculated nor infected. Care was taken that the infecting virus was given in equal doses, the skeptical Colin shaking the phial himself and the injections being alternated between protected and unprotected animals at the suggestion of other critical observers.[79] The results could hardly have been more conclusive; of the infected animals all the uninoculated died and all the inoculated survived. The discovery was immediately hailed as a triumph and applied practically.

Pasteur's method for immunization against rabies was more difficult to validate conclusively. It was recognized that the disease was fatal once a person was infected, and Pasteur was able to show that his vaccines gave recovery in patients believed to have been infected. The claim for the effectiveness of the injections was doubtless reasonable, but it has subsequently been very difficult to appraise modifications of Pasteur's technique since treatment cannot morally be withheld though evidence of genuine infection is often scanty.[80]

Increasing knowledge of metabolism and nutrition was followed by corresponding advances in therapy. One of the earliest dietetic trials was made by Takaki in the Japanese Navy. Study of official records showed him that until 1883 beriberi accounted for one-half of the fatalities and much of the invaliding. After making a review of the possible causes in the same way as Lind did 150 years earlier, he decided that the food must be at fault. A more generous diet was introduced in 1884 and from then beriberi began to decline and was eliminated by 1886.[81] This was achieved in ignorance of vitamins and success was attributed rather to increased protein allowances.

In 1891, treatment of myxedema by thyroid extract was introduced by Murray.[82] Kocher had described the development of myxedema in man following extirpation of the thyroid gland and Horsley had produced similar results

experimentally in monkeys. Murray had some diffidence in reporting a single successful case since "some cases do improve to a certain extent when untreated," but he considered that return of perspiration and menstruation after 4 years suggested real benefit. An efficient, although perhaps not conclusive, trial was here achieved by careful observation of the effects of a simple replacement therapy and its results were fully supported by subsequent experience.

The continued rapid growth of bacteriology at the end of the nineteenth century led to the preparation of a therapeutic serum for diphtheria. Behring, Boer, and Kossel[83] described the early trials of this, quoting 30 patients treated of whom 6 died, and comparing this result with the then usual mortality of about 50 per cent and with the previous year's experience at the same hospital when 21 of 32 patients died. The authors said that more cases would be required to prove the value of the new treatment. Statistically their figures are significant if the cases were comparable throughout.

OL.
10
959

The following year the rival French school published a study of 300 diphtheria patients treated by serum.[84] A mortality of 26 per cent was observed after various corrections for anomalous cases, and this was compared with a 51 per cent mortality during the previous 4 years at the same hospital and with a mortality of 60 per cent in concurrent cases at another hospital. The series were compared as far as possible in respect to type of illness and severity and it was shown that the patients treated with serum had a lower mortality in each group.

An attempt to eliminate the uncertainities of inadequately controlled trials of this treatment was made by Fibiger[85] in Denmark. He used the serum for alternate cases of diphtheria and analyzed the "treated" and "untreated" series for comparability in age, symptoms, and severity. He then compared the results, within age and severity groups, for mortality, extent of diphtheritic membrane, pyrexia, paralysis, and albuminuria. He was able to show that the mortality was less among the serum-treated, that the membrane disappeared sooner, and that pyrexia, albuminuria, and paralyses were unaltered. This excellently conducted trial was less valuable than it otherwise might have been because, although adequate numbers were studied, the disease at the time happened to be very mild—the over-all mortality in 488 cases studied was only 8 per cent.

Other studies of immunity at the end of the nineteenth century led to the introduction of prophylactic inoculation for typhoid fever. Wright[86] reported the comparative statistics of incidence of the disease in various units of the Army in India. Of more than 11,000 men about 3,000 had received inoculation; less than 1 per cent of these developed typhoid. Among the remaining 8,000 unprotected men the incidence was 2.5 per cent. The circumstances of this trial prevented randomization of the prophylaxis among the men and the possibility of differerences in susceptibility and exposure cannot therefore be excluded. A more evident difficulty arose from the necessary reliance upon many different observers to diagnose the disease and record its presence or absence. These observers were army doctors who would inevitably vary in skill and in interest in the trial. While the results clearly suggest a degree of protection achieved by the inoculation, they cannot be considered conclusive.

THE TWENTIETH CENTURY

The enormous development of organic chemistry, parallel with the previously mentioned growth of bacteriology, laid the foundation for chemotherapy, the greatest contribution of our age to medical treatment. At the turn of the century Ehrlich began his search for trypanocidal agents. His success with arsphenamine in vitro led to its trial first for syphilis in the rabbit, then for the human disease. From the animal experiments Ehrlich believed he had achieved his aim of a "magna therapia sterilans" and the early trials of arsphenamine were made on this assumption. Wechselmann[87] reported on 9 patients, each treated with a single dose of the new drug. In several, rapid clinical improvement occurred but failure in others was attributed to the use of too small a dose. Although the Wassermann test had been introduced 4 years earlier, it does not seem to have been used routinely in these early trials of "606." Relapses soon showed that persistent treatment was necessary, and, as further chemotherapeutic substances were developed, increasingly stringent tests of their effectiveness were used. Such criteria are listed by Moore and others[88] and include the rate of disappearance of surface organisms, of clinical healing of lesions, and of seroreversal; the proportion of clinical and serologic relapses; the occurrence of spinal fluid changes; and the number of resistant cases.

Meanwhile, several established drugs were subjected to renewed trial. Digitalis, which had been studied so profitably by Withering in the eighteenth century, was assessed in more modern terms by Mackenzie.[89] The general approach was almost unchanged; a wide range of carefully studied cases was reviewed in which standard preparations of the drug had been used along with the latest techniques of examination. The results confirmed and elaborated Withering's findings.

A trial of emetine by injection for the treatment of amebic infections was made by Rogers,[90] who published a favorable report based on 3 patients who were not able to take ipecachuana by mouth. All made quick recoveries on a low dosage. The small number of patients, the vagaries of the usual clinical course, and the absence of a long follow-up make this trial by itself unconvincing, but other workers soon confirmed the results.

World War I raised new problems of large scale prophylaxis and treatment in medicine and surgery. Routine tetanus antiserum for the wounded was introduced at the end of 1914. The incidence and severity of tetanus fell dramatically,[91] though it is conceivable that change of terrain or of surgical procedures may have contributed to this result. The evidence in favor of prophylactic antityphoid and anticholera inoculation and the value and limitations of the chi square method of testing such data were reviewed by Greenwood and Yule.[92] In the treatment of wound shock the use of gum-saline solutions was a major advance.[93] This was based on sound physiologic theory and its value was first demonstrated upon experimental shock in animals. Its freedom from immediate toxicity was also tested and its effects in the resuscitation of a series of patients were carefully observed;[94] the delayed effects upon the liver which have since caused gum-saline to be abandoned were overlooked.

VOL
10
195

The aftermath of the war provided both the stimulus and the opportunity for accurate clinical trials of the treatment of rickets. In Vienna the disease became very common in children's hospitals and was there believed to be of infective origin. The diet given in treatment was grossly deficient in vitamins. A team of British research workers was thus able to make a conclusive controlled trial of vitamin supplements and irradiation, with corrections for other factors such as season of year and exposure to sunlight.[95] An exceptional type of control was available for some of the observations. A child treated for scurvy and rickets was found to have a twin brother elsewhere in the hospital with whom progress could be compared. Photographs of the two show the enormous improvement following vitamin therapy.[96] Another problem repeatedly attacked was the assessment of prophylaxis against the common cold and other respiratory infections by immunization.[97,98] These trials, though having unspectacular results, developed controlled investigations upon large numbers of persons to a new stage. For instance in the second study the element of suggestion was minimized by treating all the subjects similarly so that only the research workers knew which had received the saline and which the vaccine injections. The first part of this trial included only volunteers, but in the second part, in order to increase the numbers of observations, nonvolunteers acted as controls. It is interesting that the two parts gave similar negative results so that in this particular example of a trial upon a fairly homogeneous student population the less stringent comparison between treated volunteers and control nonvolunteers seems to have been sufficient.

OL.
10
959

The two great therapeutic advances of the postwar decade were of a physiologic type. Banting and Best's introduction of insulin for the treatment of diabetes was initially based upon experiments on pancreatectomized dogs;[99] careful trial followed of the effects upon the blood sugar, glycosuria, ketonuria, and general condition of diabetic patients.[100] The other outstanding development was the demonstration of the value of liver treatment for pernicious anemia by Minot and Murphy.[101] These workers had previous wide experience with the disease and knew that they might expect spontaneous remissions in about one-third of cases. They gave the liver diet to 45 patients in relapse and found that all improved and remained cured so long as they continued treatment; 3 who stopped treatment suffered relapse. Hematologic studies showed that the first response was the rise in reticulocytes and that clinical improvement could be recognized a few days later. This trial shows how much can be achieved without formal control cases when a markedly successful treatment is tried upon a chronic condition and the results are cautiously interpreted.

The treatment of lobar pneumonia by sera was the subject of several extensive and careful trials. Cecil[102] treated several hundred cases in certain wards with serum and in others without. Typing of the pneumococci made precise comparisons possible and the mortality results showed an improvement with the serum in all but the Type III infections which were almost unchanged. Park, Bullowa, and Rosenblueth[103] made a similar trial treating alternate patients with serum. Typing was also used but the authors point out that randomization of cases within serologic types was not feasible since treatment must begin before

a typing result is available. The conclusions were very similar to those of Cecil and were subsequently confirmed in other trials.

New assessments of drugs used for chronic cardiac disease were made by Sir Thomas Lewis and his school. Harris[104] made a study of the long-term results of quinidine treatment of auricular fibrillation. Lewis[105] gave a detailed clinical description of the effect of amyl nitrite upon 4 cases of angina pectoris. A complementary study of the same and alternative drugs was made by Wayne and Laplace[106] upon the functional capacity of patients with angina estimated by the amount of exercise required to precipitate pain. This group of trials had in common an emphasis upon precise diagnosis and careful clinical observation of relatively few patients. The diseases studied were of chronic and sustained type so that the treated patients themselves could be considered as controls. The careful accessory observations provided not only an estimate of the utility of the remedies but also new information on their mode of action.

The hypothesis, based on animal experiments, that vitamins had a prophylactic value against infection was tested by Green, Pinder, Davis, and Mellanby.[107] They studied the incidence of puerperal sepsis in 550 women when supplementary vitamins A and D were given to alternate admissions. The treated and control series were carefully analyzed for comparability in respect to parous state, age, and the like. The results showed a consistent improvement in those receiving the vitamins in all clinical severities of sepsis. The official standard of puerperal morbidity was felt to be inadequate since, in spite of the evident and statistically significant lowered incidence of sepsis, judged by the morbidity standard, there was no significant advantage.

The years since 1935 have possibly seen more clinical trials than occurred in the whole of previous medical history. It is evidently not feasible to review even briefly all these recent therapeutic studies. Instead a few examples only will be considered representing the fields of chemotherapy, substances used for their physiologic effects, and antibiotics.

Domagk's announcement in 1935[108] of successful treatment by Prontosil of experimental streptococcal infection in animals led to numerous trials of the sulphonamides for human disease. Colebrook and Kenny[109] and Colebrook and Purdie[110] showed that Prontosil and sulfanilamide were effective in puerperal sepsis. They observed a mortality rate of 8 per cent which they compared with the finding of 22 per cent for the previous years at the same hospital where the condition had long been studied. In 1937 Snodgrass and Anderson[111] showed in a controlled trial upon 312 patients that Prontosil shortened the durations of spread, of primary pyrexia, and of toxemia in erysipelas. Control groups received ultraviolet radiation therapy or scarlet fever antitoxin. Their comparability with the group receiving Prontosil was checked in respect to age and state and severity of the disease. The authors noted that mortality in this condition was too small for it to be a useful criterion of comparison. Evans and Gaisford[112] demonstrated the value of sulfapyridine for lobar pneumonia by a controlled trial in which 100 alternate patients received the drug while the other 100 had routine nonspecific therapy in the wards of colleagues. The comparability of the patients in sex and age distribution were tested. The case mortality of the

sulfapyridine series was 8 per cent compared with 22 per cent in the controls. Further comparisons are offered with the results at other hospitals and previously at the same hospital which showed a mortality similar to that of the control series in the trial. Banks[113,114] showed that the mortality from meningococcal meningitis could be reduced spectacularly. He first used a variety of doses combined with serum treatment and in this series 16 per cent died; those receiving larger doses did better, so in a second series adequate amounts of sulfanilamide or sulfapyridine only were given and of 67 patients so treated only one died. The trial also provided valuable information on dosage, route of administration, and the resulting levels in blood and cerebrospinal fluid. No formal controls were arranged, comparisons being implied with the previously uniformly high mortality in this clearly defined disease and with the intermediate results in the first inadequately treated cases. A trial of sulfonamides by Wagle and others[115] during an epidemic of plague in Bihar is an interesting example of what can be achieved in difficult circumstances. Successive admissions to an emergency center were treated either with antiplague serum or iodine, a common Indian remedy, or sulfonamide. Treatment was commenced before completion of diagnosis by culture; negative cases were subsequently excluded. Other exclusions had to be made because some patients discharged themselves, but enough cases remained to show that sulfonamides reduced the mortality to less than half that of the iodine-treated controls. A similar advantage was also shown when the more severe septicemic cases were considered alone.

Another field of advance in chemotherapy has been that of the antimalarial drugs. Methods of study analogous to those of the sulfonamide trials have been used but there have been further special features. Therapeutic infections with malaria in the treatment of general paralysis provided a supply of known cases of standard type on which antimalarial therapy could be tested. The results of such studies were reported by the League of Nations Malaria Commission.[116,117] The definite course and symptomatology of the induced disease and the clear criteria of cure made it possible to give precise comparisons of the effects of different drugs and treatment routines. The ease with which the experimental disease could be produced made it possible also to use otherwise normal volunteers for trials either in hospitals where detailed responses could be studied or under various conditions of stress simulating active service conditions. The latter method was used in a large scale wartime trial on the chemotherapeutic suppression and prophylaxis of the disease in the southwest Pacific area.[118,119] The research team included entomologists to supply the infected mosquitoes, pathologists to study the effect of the drugs upon parasites and patients, and clinicians to look after the patients, supervise therapy, and study the clinical course of the disease. Infected treated volunteers were compared with infected untreated controls for occurrence of the disease, appearance of parasites in the blood, and capacity to infect other volunteers by subinoculation. Not only was the great value of mepacrine and later that of Paludrine clearly established, but new ancillary information on the life cycle of the malarial parasite was obtained; further, this and the prewar trials at last firmly assessed the antimalarial status of quinine after it has been in clinical use for 300 years.

In addition to the spectacular developments of chemotherapy for infection there has been progress in the use of substances acting upon the disordered physiology of disease. Transfusion of burned patients with plasma is an example. Black[120] made a careful study of the biochemical and blood volume changes produced by such treatment in a small series of patients. The findings before, during, and after the transfusion were compared and further comparisons were made between patients receiving different quantities of plasma for different severities of burning. A therapy with a less definite rationale but which presumably acts upon the physiology of the circulation is the treatment of migraine with ergotamine tartrate. Lennox and von Storch[121] reported a trial upon 120 patients in whom other treatment had failed. Given by various routes, the drug relieved consistently the attacks of 109 of these people; this was contrasted with the previous unsuccessful treatment of this chronically recurrent condition in the same patients. O'Sullivan[122] reported a similar trial; experience over several years had suggested to him the value of ergotamine, so he made a trial of it on 97 patients during 2 years. This showed abrupt termination of attacks in 89 of them and of 1,132 headaches 92 per cent were immediately relieved after the use of the drug. No data are given to indicate how many attacks would have subsided spontaneously nor in how many the psychologic effects of the therapy of this largely subjective disease may have played a part. Comparisons are available only with the previous general experience of the disease, the observer's previous experience of the disease when treated with other drugs, and the previous course in the same patients. The striking results give presumptive evidence of the value of ergotamine but cannot be considered conclusive.

The possible value of methionine for infective hepatitis was suggested by biochemical and histologic studies in animals. Wilson, Pollock, and Harris[123] studied 100 cases of this disease, giving alternate patients a dose of methionine in orangeade, while controls received the orangeade only. Treatment was otherwise similar and the two series were checked for comparability in respect to age, stage and severity of disease. Comparisons of the durations of anorexia, jaundice, liver enlargement, liver tenderness, and hospital treatment and of the incidence of relapses showed slight and insignificant advantages among the patients treated with methionine. Higgins and co-workers[124] made a similar trial upon a smaller series of alternate cases. They followed these patients very thoroughly and checked the quantities of extraneous methionine given in the diet. Their conclusions were similar to those of the other workers.

Penicillin, still the outstanding antibiotic, was first subjected to full clinical trial by Abraham and associates.[125] The scarcity of the material determined the selection of a few desperately ill patients, several of them children, so that maximum results could be obtained with small quantities. Dramatic improvement resulted although administration was stopped early and some relapses then occurred. Apart from trouble with pyrogenicity due to impurities, the low toxicity was established. When more penicillin became available the trial was extended on similar lines to more cases and systemic dosage was arranged to maintain inhibitory blood levels against the causal organisms.[126] The development of methods of detection and estimation of the drug greatly aided the pre-

VOL. 10 1959

cision of this trial and its conclusions have since been amply confirmed. No formal controls were used but the implied comparison of the clinical course of these patients under treatment with that before treatment and with the known prognosis is sufficiently convincing. Further confirmation was provided by the demonstration of the mode of action of the drug in that bactericidal levels could be shown in body fluids which corresponded to changed bacteriologic findings and clinical improvement. The effects of local application of penicillin in 15 patients with war wounds were investigated by Pulvertaft[127] who reported a great improvement in infections with gram-positive organisms but no effects upon gram-negative infections. The clinical effects of penicillin in the treatment of such wounds were studied in North Africa and in an extensive trial conducted in the 21 Army Group.[128] For this the cooperation of surgeons was enlisted to compare the effects of penicillin with those of the most favored alternative in the treatment of similar conditions of infection. The period of healing was used as criterion of success. It was soon found that allocation of cases was not impartial since surgeons did not feel justified in withholding penicillin from the more seriously injured. In spite of this adverse loading of the scales the penicillin group did demonstrably better. In retrospect this trial would appear to have been more successful than might have been anticipated with such indirect organization and multiple observers. Had penicillin been less effective the biased control might have caused an inconclusive result; since the effect was so great perhaps a smaller and more precise trial would have demonstrated it with greater efficiency. While supplies of penicillin were still short, a large-scale planned trial of its value in subacute bacterial endocarditis was begun.[129,130] Fourteen treatment centers were organized, and different treatment schedules allotted to different centers. No controls were arranged, since the universally bad prognosis of untreated cases provided sufficient comparison. All centers at first gave a total of 5 million units, but the length of the course was varied. This showed that a prolonged course gave the highest rate of cure for a given total dose, so in the second period all patients had a long course but the daily dosage was varied. By this method of successive approximation invaluable information was obtained on the treatment of this extremely serious disease. As it happened, the optimal dosage does not seem to have been "straddled" in either of the sections of the trial and work has continued on yet higher doses for longer periods. A less severe condition in which the relatively slight effect of penicillin makes precise controls imperative is that of finger pulp infections. Various series have been published in which the results with use of penicillin have been compared with those of preceding or subsequent cases. Conflicting conclusions were drawn but a carefully controlled trial clarified the problem.[131] Patients were first admitted to the experimental series and were then randomly allocated to the different treatment groups. Detailed clinical progress was recorded and a considerable advantage could be demonstrated in patients receiving systemic penicillin. As frequently happens in such studies other improvements in treatment were developed during the trial but each such improvement was applied equally to "treated" and control cases. Had concurrent controls not been used, the effect of these changes would have been confused with that of the treatment under

trial. There would seem to be many opportunities for trials of this type where several factors in the therapy, such as incision, rest, and chemotherapy, are important but where the optimal balance cannot easily be determined.

## CONCLUSIONS

Although the clinical trial is one aspect of scientific medicine, its status as a scientific experiment is modest. The full sequence of observation-theory-hypothesis-experiment-validation is not usually invoked. The "hypothesis" is usually very simple—that the test treatment is better than some other treatment. Procedurally this is equivalent to problems in technology, such as the building materials to be preferred for certain purposes or the most efficient driving mechanism for a piece of machinery. Such problems were being solved by a simple form of experimentation long before the modern scientific era. Their solution does not necessarily imply any appreciation of inductive reasoning such as was introduced by Francis Bacon and Galileo in their advancement of scientific method. It is therefore not surprising to find that the principles on which satisfactory trials could be based were already implied in classical writings,[9,11] described in detail by Avicenna,[15] and the peculiar difficulties of experiment in medicine recognized by Roger Bacon.[17] In spite of this there are very few records of tests of therapy until the last two centuries. Why this should be must remain speculative, but certain factors can be seen to be relevant:

1. *Reverence for Authority.*—Until modern times the current habit of mind was to deduce "facts" by argument from premises approved by tradition and authority without appeal to experimental validation. Even when observation ran counter to such "facts," as for instance in the case of Jesuits' bark, it was still felt that perhaps in some more mysterious way authority must still be right. Many of the leaders of scientific development, such as Roger Bacon, Leonardo da Vinci, and Harvey, continued to favor traditional therapy, and even pioneers of clinical trials often failed to rely upon their own findings. Having demonstrated the effect of fruit in the cure of scurvy, Lind recommended change of air as the first requirement of treatment; similarly after Louis had shown the futility of bleeding in a variety of conditions he still continued to advise it for the hemoptysis and pleurisy of phthisis. The emergence of the modern therapeutic trial is thus an aspect of the growth of reliance upon impartial observation as the foundation of scientific medicine.

2. *The Relationship of Doctor and Patient.*—Roger Bacon early called attention to this difficulty in the scientific approach to medicine. The patient needs to be cured, and there are important ethical considerations in an experimental therapy; the doctor tends to adopt an authoritarian attitude; the organization of careful comparative study is inconvenient in a busy practice.

3. *The Paucity of Records.*—The recording and publishing of scientific evidence is a relatively recent development. Printing was a prerequisite for large-scale dissemination of new findings and even after the development of printing it only gradually became a recognized and easy method of scientific communication.

VOL. 10 1959

4. *Lack of Facilities for Investigation.*—Until the eighteenth century, hospitals were inadequate and unsuitable for sustained studies. Techniques of accurate diagnosis and assessment of cure were lacking, and until hospitals had been organized as functioning units there were few opportunities for specialization so that sufficient numbers of cases of the same type could be studied by one person.

5. *Polypharmacy.*—From ancient times through the Middle Ages and even now in some places, polypharmacy pervades medical treatment. It implied exact information on the effect of the ingredients of a mixture and of their interaction when in combination. This information did not really exist, and the practice of polypharmacy helped to perpetuate this ignorance.

6. *Lack of Active Remedies.*—This must be a main reason for the absence of early clinical trials. Apart from a few agents such as mercury, quinine, opium, and purgatives, few really active materials were available until the nineteenth century, when with the development of organic chemistry and later biochemistry and bacteriology a multitude of new and active products stimulated the demand for reliable clinical trials. Parallel with the discovery of new remedies came a more critical appraisal of current treatment, and it is noteworthy that many of the early trials discredited the claims of charlatans or of traditional therapy.

The history of clinical trials gives an indication of the circumstances in which they can be successfully and easily performed, and conversely the sources of difficulty. The therapy should be simple and measurable so that it can be applied similarly on all occasions. The condition treated should be clearly definable and either of regular severity in different patients or of measurable severity so that similar cases can be grouped together. This can sometimes be achieved by experimental induction of the disease. A satisfactory comparison must be available between treated cases and cases otherwise similar but not receiving the treatment under test. This may be provided by a known uniform prognosis in untreated cases, by use of alternating control periods in chronic diseases, or by randomization of different cases into treated and untreated series. The criteria of success and failure should be clear and should be recognizable early. Preferably they should be chosen so that the control cases show a high proportion of "failure" since the trial may otherwise be prolonged or inconclusive. The trial will further increase in efficiency the more regular is the clinical material and the more constant and marked the effect of the therapy. A controlled trial offers an enormous increase of precision over other methods, but it is still imperative that the soundness of the comparison be critically examined. Results must be carefully interpreted so that conclusions drawn are within the limits of the logical structure of the trial and due regard to special experimental conditions should be paid when the findings are applied clinically.

A difference between trials in medicine and surgery has been noted: whereas surgery made great progress in ancient times, the development of medical therapy was delayed until recent years. Part of this discrepancy can be attributed to the differences in the respective clinical trials demanded. Much of surgery deals with lesions of simple etiology and diagnosis, and the application of simple mechanical principles often produces immediate spectacular results. Suturing

of wounds and reduction of dislocations and fractures give evidence of success at once and leave little opportunity for elaborate theorizing. The fundamentals of such treatment had already been worked out in ancient Greece and Egypt. By contrast, few medical conditions could be understood until ancillary scientific methods had been developed. The aspects of medical diseases which could be appreciated, such as fever, dropsy, jaundice, or wasting, were of such varied etiology that any but the simplest treatment by rest or diet would have inconstant effects and even if active agents were available the results would seldom be immediately apparent. It is thus not surprising that superstition and scholastic theorizing thrived and few therapeutic advances were made.

It seems possible that the success of simple surgical trials in the past has led workers to believe that similar methods remain applicable to the complex procedures of modern surgery. In his criticism of the validation of gastrojejunostomy, Ryle[132] pointed out that trial and error methods have been repeatedly misleading and that, in the absence of controls, adequate follow-up of patients, and statistical analysis, such an operation cannot be fairly assessed. The effects of extensive surgery can be quite as complex as those of elaborate medical therapy. Not only must immediate operative mortality and the achievement of anatomic success be considered but also late effects of altered function and possibly secondary disease. Such a situation demands the same care in the planning of a satisfactory clinical trial as does the detailed assessment of a new and potentially dangerous antibiotic.

The unintentional "trials" of wound treatment made by Paré and Hunter were incidental to the usual routine of practice. In a comparable way some modern trials have been incidental to other studies. The observations of Anning[133] on the effects of calciferol upon chilblains provide an example. The occurrence of chilblains in patients receiving this treatment for other complaints was compared with their occurrence in patients not so treated. No difference was found between the groups. Studies on the pathogenicity of certain organisms can also be profitably combined with tests of antibacterial therapy; this has been done in trials of local antibiotics for burns.[134,135] This integration of therapeutic trials with other aspects of medicine is further shown in the diagnostic advances made in Withering's study of digitalis and the Vienna studies of rickets and in the new information on the life cycle of *Plasmodium vivax* which resulted from Fairley's trials of mepacrine and Paludrine. A thorough trial inevitably involves a close study of patients, often a large number of similar ones. It is therefore not surprising that other discoveries may be made even though they are not anticipated.

The present century has seen an intensification of therapeutic problems, not least by two World Wars, an increase of facilities for their study in hospital and laboratory accommodation and in research organizations. As well as the continued production of new remedies from bacteriology, there have been added contributions from organic chemistry such as the arsenicals and later the sulfonamides. More recently still, from bacteriology, the antibiotics have been introduced and some of them have been synthesized. There is increasing appreciation of the importance of adequate controls and in other fields new sta-

tistical methods have been developed, many of which still wait to be fully applied to clinical trials.

The stage is set for great therapeutic advances in which clinical trials must play an important part. This is not to say that all features of therapy should be subjected to trial; some are too minor for the effort and expense to be justifiable, others too complex to be feasible. A reasonable balance would seem to be the establishment of crucial principles by careful clinical trials followed by practical application of the conclusions, modified where necessary in the light of scientific and clinical knowledge. At a later stage an assessment of results may be possible which will check whether or not the modifications are justifiable. Further exploratory tests can be made as new treatments become available or inconsistencies appear in results, and in due course a new full-scale trial will be required. In this way therapeutics should advance by a series of steps, careful trials providing at each stage a firm basis for exploratory hypotheses and experiments which can assist in determining the next advance.

OL.
10
959

REFERENCES

1. British Museum: Introductory Guide to the Egyptian Collections, London, 1930, British Museum, pp. 71, 98.
2. Bryan, C. P.: The Papyrus Ebers, London, 1930, Geoffrey Bles. Ltd.
3. Breasted, J. H.: The Edwin Smith Surgical Papyrus, Chicago, 1930, University of Chicago Press.
4. Ranke, H.: Medicine and Surgery in Ancient Egypt, Bull. Inst. Hist. Med. **1**:237, 1933.
5. Dawson, W. R.: The Beginnings: Egypt and Assyria, New York, 1930, Paul B. Hoeber, Inc.
6. Jastrow, M.: Babylonian-Assyrian Medicine, Ann. M. Hist. **1**:231, 1917.
7. Gordon, B. L.: Medicine Among the Ancient Hebrews, Ann. M. Hist. **4**:219, 1942.
8. Hippocrates: Regimen in Acute Diseases. *In* Works, translated by W. H. S. Jones, London, 1923, Wm. Heinemann, Vol. ii, p. 63.
9. Hippocrates: Precepts. *In* Works, translated by W. H. S. Jones, London, 1923, Wm. Heinemann, Vol. i, p. 313.
10. Singer, C.: From Magic to Science, London, 1928, Benn, p. 59.
11. Celsus, A. C.: De Medicina, translated by W. G. Spencer, London, 1935, Wm. Heinesmann.
12. Gunther, R. T.: The Greek Herbal of Dioscorides, Oxford, 1933, Magdalen College, Oxford University.
13. Garrison, F. H.: An Introduction to the History of Medicine, ed. 3, Philadelphia, 1922, W. B. Saunders Co.
14. Mettler, C. C.: History of Medicine, Toronto, 1947, The Blakiston Company, p. 190.
15. Crombie, A. C.: *In* Avicenna, Scientist and Philosopher, edited by G. M. Wickens, London, 1952, Luzac & Company, Ltd., p. 89.
16. Harington, J.: The School of Salernum; Regimen sanitatis salernitanum, 1607 (reprinted, 1920, New York).
17. Bacon, R.: De erroribus medicorum, translated by E. T. Withington. *In* Essays on the History of Medicine Presented to Karl Sudhoff, London, 1924, Oxford University Press, p. 139, c. 1268.
18. Allbutt, T. C.: Science and Mediaeval Thought, London, 1901, C. J. Clay & Sons.
19. Lilley S.: Men, Machines and History, London, 1948.
20. Nicaise, E. (tr.): Chirurgie de Maitre Henri de Mondeville, Paris, 1893, J. B. Baillière et fils.
21. Cumston, C. G.: An Introduction to the History of Medicine, London, 1926, Kegan Paul, French, Trubner, & Co., Ltd., p. 372.
22. Paracelsus: Selected Writings, edited by J. Jacobi, translated by N. Guterman, London, 1951, Routledge & Kegan Paul, Ltd., p. 126.
23. Sherrington, C. S.: The Endeavour of Jean Fernel, London, 1946, Cambridge University Press.
24. Richter, J. P.: Leonardo da Vinci, Oxford, 1939, Oxford University Press, vol. ii, p. 239.
25. Lilley, S.: Cause and Effect in the History of Science, Centaurus **3**:58, 1953.
26. MacCurdy, E.: The Notebooks of Leonardo da Vinci, London, 1938, Jonathan Cape.

27. Packard, F. R.: The Life and Times of Ambroïse Paré, ed. 2, New York, 1925, Paul B. Hoeber, Inc., pp. 27, 163.
28. Gibson, R. J. H.: Outlines of the History of Botany, London, 1919, p. 14.
29. Drummond, J. C., and Wilbraham, A.: The Englishman's Food, London, 1939, Jonathan Cape.
30. Bacon, F.: De augmentis scientiarum. *In* Works (1818), London, new ed., 1623, M. Jones, vol. vi, p. 153.
31. Harvey, W.: De motu cordis, Frankfurt, 1628; new edition by Keynes, G., London, 1928, Cambridge University Press.
32. Shryock, R. H.: The Development of Modern Medicine, London, 1948, V. Gollancz, p. 24.
33. Boyle, R.: The Usefulness of Natural Philosophy, Oxford, 1663. *In* Works, edited by Birch, 1744, Vol. i, p. 505.
34. Brown, H.: Jean Denis and Transfusion of Blood, Isis **39**:15, 1948.
35. Haggis, A. W.: Fundamental Errors in the Early History of Cinchona, Bull. Inst. Hist. Med. **10**:417, 1941.
36. Scott, H. H.: A History of Tropical Medicine, London, 1939, Edward Arnold & Co., p. 223.
37. Payne, J. F.: Thomas Sydenham, London, 1900, F. Fisher Unwin, p. 152.
38. Osler, W.: John Locke as a Physician, Lancet **2**:1115, 1900.
39. Haggard, H. W.: The Lame, the Halt and the Blind, London, 1932, Wm. Heinemann, p. 35.
40. Wiseman, R.: Severall Chirurgicall Treatises, London, 1676, R. Royston, p. 123.
41. Maitland, C.: Account of Inoculating the Small Pox, London, 1722, J. Roberts.
42. Creighton, C.: A History of Epidemics in Britain, Cambridge, 1894, University Press, Vol. ii, p. 725.
43. Dover, T.: The Ancient Physician's Legacy to his Country, London, 1733, A. Bettesworth & C. Hitch.
44. Bradley, H.: Physical and Philosophical Remarks on Dr. Dover's Late Pamphlet, London, 1733, C. Rivington.
45. Hales, S.: An Account of Some Experiments and Observations on Mrs. Stephens' Medicines for Dissolving the Stone, London, 1740, J. Woodward.
46. Heberden, W.: Antithriaka, an Essay on Mithridatium and Theriaca, London, 1745.
47. Berkeley, G.: Siris; a Chain of Philosophical Reflexions and Inquiries Concerning the Virtues of Tar Water, London, 1744, W. Innys.
48. Clark-Kennedy, A. E.: Stephen Hales, London, 1929, Cambridge University Press.
49. Hales, S.: An Account of Some Experiments and Observations on Tar Water, London, 1747, R. Manby & H. S. Cox.
50. Lind, J.: A Treatise of the Scurvy, Edinburgh, 1753, Sands, Murray, & Cochran.
51. Cook, J.: The Method Taken for Preserving the Health of the Crew of His Majesty's Ship *The Resolution* During Her Late Voyage Round the World, Philos. Trans. **66**:402, 1776.
52. Franklin, B.: Letter to John Pringle. *In* Benjamin Franklin's Experiments, edited by I. B. Cohen, (1941) Harvard, 1757.
53. Haygarth, J.: Of the Imagination as a Cause and as a Cure of Disorders of the Body, new ed., Bath, 1801, R. Crutwell, p. 3.
54. Withering, W.: An Account of the Foxglove and Some of Its Medical Uses With Practical Remarks on Dropsy and Other Diseases, Birmingham, 1785, G. C. J. & J. Robinson.
55. Roddis, L. H.: William Withering and the Introduction of Digitalis Into Medical Practice, Ann. M. Hist. **8**:93 and 185, 1936.
56. Pearson, H.: Doctor Darwin, London, 1930, J. M. Dent.
57. Fowler, T.: Medical Reports of the Effects of Arsenic, 1786, J. Johnson.
58. Hunter, J.: A Treatise on the Blood, Inflammation and Gun-Shot Wounds, London, 1794, G. Nicoll, p. 531.
59. Jenner, E.: An Inquiry Into the Causes and Effects of the Variolae Vaccinae, London, 1798, S. Low.
60. Pearson, G.: An Inquiry Concerning the History of Cowpox, London, 1798, J. Johnson.
61. Woodville, W.: Reports of a Series of Inoculations for the Variolae Vaccinae or Cowpox, London, 1799, J. Phillips & Sons.
62. Waterhouse, B.: A Prospect of Exterminating the Small-Pox, Boston, 1800, Cambridge.
63. Pearson, J.: Observations on the Effects of Various Articles of the Materia Medica in the Cure of Lues Venerea, Illustrated with cases, London, 1800, J. Callow, Introd. p. 1.
64. Cobbett, W.: The Rush Light, New York, 1800, Wm. Cobbett.
65. Laplace, P-S.: Théorie Analytique des Probabilités, Paris, 1812.
66. Louis, P. C. A.: Researches on Phthisis, ed. 2, translated by W. H. Walshe, London, 1843, Sydenham Society, pp. 551, 552.
67. Louis, P. C. A.: Essay on Clinical Instruction, translated by P. Martin, London, 1834, S. Highley, pp. 26, 27, 28.
68. Louis, P. C. A.: Recherches sur les effets de la saignée, Paris, 1835, De Mignaret.

VOL.
10
1959

69. Magendie, F.: Formulaire pour la préparation et l'emploi de plusieurs nouveaux médicamens, ed. 3, Paris, 1822, Méquignon-Marvis, preface, p. vi.
70. Bartlett, E.: An Essay on the Philosophy of Medical Science, Philadelphia, 1844, Lea & Blanchard.
71. Osler, W.: An Alabama Student, and Other Biographical Essays, Oxford, 1908.
72. Holmes, O. W.: Currents and Countercurrents in Medical Science. *In* Works, 1891, Vol. ix, p. 185.
73. Bennett, J. H.: The Restorative Treatment of Pneumonia, Edinburgh, 1865, A. C. Black.
74. Sutton, H. G.: Cases of Rheumatic Fever Treated for the Most Part by Mint Water, Guy's Hosp. Rep. **2**:392, 1865.
75. Duncum, B. M.: The Development of Inhalation Anaesthesia, London, 1947, Oxford University Press.
76. Snow, J.: On Chloroform and Other Anaesthetics, London, 1858, J. Churchill.
77. Lister, J.: On the Effects of the Antiseptic System Upon the Salubrity of a Surgical Hospital, Lancet **1**:4 and 40, 1870.
78. Times: Latest Intelligence, France, London, June 3, 1881, p. 5.
79. Vallery-Radot, R.: Pasteur, 1822-1895, Paris, 1922.
80. Harvey, W. F., and Acton, H. W.: An Examination Into the Degree of Efficicacy of Anti-rabic Treatment, Ind. J. M. Res. **9**:852, 1922.
81. Takaki, K.: On the Preservation of Health Among the Personnel of the Japanese Navy and Army, Lancet **1**:1369, 1906.
82. Murray, G. M.: Note on the Treatment of Myxoedema by Hypodermic Injections of an Extract of the Thyroid Gland of a Sheep, Brit. M. J. **2**:796, 1891.
83. Behring, E., Boer, O., and Kossell, H.: Zur Behandlung diphtheriekranker Menschen mit Diphtherieheilserum, Deutsche med. Wchnschr. **19**:389 and 415, 1893.
84. Roux, E., Martin, L., and Chaillou, A.: Trois cent cas de diphthérie traité par le serum antidiphthérique, Ann. Inst. Pasteur **8**:640, 1894.
85. Fibiger, J.: Om Serumbehandlung af Difteri, Hospitalstidende **4**:6, 309 and 337, 1898.
86. Wright, A. E.: Remarks on the Results Which Have Been Obtained by the Anti-typhoid Inoculations, Brit. M. J. **1**:122, 1900.
87. Wechselmann, W.: Über die Behandlung der Syphilis mit Dioxydiamido-arsenobenzol, Berl. klin. Wchnschr. **47**:1261, 1910.
88. Moore, J. E., Stanton, M. H., Robinson, H. M., and Eagle, H.: Response of Quantitatively Titred Wassermann Test in Early Syphilis to Treatment With Five Different Arsenical Drugs, Am. J. Syph. **20**:503, 1936.
89. Mackenzie, J.: Diseases of the Heart, London, 1908, Hodder & Stoughton, pp. 269, 334.
90. Rogers, L.: The Rapid Cure of Amoebic Dysentery and Hepatitis by Hypodermic Injections of Soluble Salts of Emetine, Brit. M. J. **1**:1424, 1912.
91. Andrewes, F. W.: *In* History of the Great War, Medical Services, Surgery of the War, London, 1922, His Majesty Stationery Office, Vol. i, p. 151.
92. Greenwood, M., and Yule, G. U.: The Statistics of Anti-typhoid and Anti-cholera Inoculations and the Interpretation of Such Statistics in General, Proc. Roy. Soc. Med. **8**:113, 1915.
93. Bayliss, W. M.: Intravenous Injections to Replace Blood. *In* Wound Shock and Haemorrhage, M. Res. Coun. Spec. Rep. Ser. No. 25, London, 1919, p. 11.
94. Drummond, H., and Taylor, E. S.: The Use of Intravenous Injections of Gum Acacia in Surgical Shock. *In* M. Res. Coun. Spec. Rep. Ser. No. 25, London, 1919, p. 135.
95. Chick, H., Dalyell, E. J., Hume, E. M., Mackay, H. M. M., Smith, H. H., and Wimberger, H.: Studies of Rickets in Vienna 1919-1922, M. Res. Coun. Spec. Rep. Ser. No. 77, London, 1923, p. 19.
96. Chick, H., and Dalyell, E. J.: Observations on the Influence of Foods in Stimulating Development in Backward Children, Brit. M. J. **2**:1061, 1921.
97. von Sholly, A. I., and Park, W. H.: Trial of Mixed Vaccine Prophylaxis Against Acute Respiratory Infections, J. Immunol. **6**:103, 1921.
98. Ferguson, F. R., Davey, A. F. C., and Topley, W. W. C.: The Value of Mixed Vaccines in Prevention of the Common Cold, J. Hyg. **26**:98, 1927.
99. Banting, F. G., and Best, C. H.: The Internal Secretion of the Pancreas, J. Lab. & Clin. Med. **7**:251, 1922.
100. Banting, F. G., Best, C. H., Collip, J. B., Campbell, W. R., and Fletcher, A. A.: Pancreatic Extracts in the Treatment of Diabetes Mellitus, Canad. M.A.J. **12**:141, 1922.
101. Minot, G. R., and Murphy, W. P.: Treatment of Pernicious Anaemia by a Special Diet, J.A.M.A. **87**:470, 1926.
102. Cecil, R. L.: Specific Treatment of Lobar Pneumonia, Arch. Int. Med. **41**:295, 1928.
103. Park, W. H., Bullowa, J. G. M., and Rosenblueth, M. B.: Treatment of Lobar Pneumonia With Refined Specific Anti-bacterial Serum, J.A.M.A. **91**:1503, 1928.
104. Harris, K. E.: Series of Cases of Auricular Fibrillation Treated With Quinidine Sulphate, Heart **14**:283, 1927.
105. Lewis, T.: Angina Pectoris Associated With High Blood Pressure and Its Relief by Amyl Nitrite, Heart **15**:305, 1930.

106. Wayne, E. J., and Laplace, L. B.: Observations on Angina of Effort, Clin. Sc. **1**:103, 1933.
107. Green, H. N., Pinder, D., Davis, G., and Mellanby, E.: Diet as a Prophylactic Agent Against Puerperal Sepsis, Brit. M. J. **2**:595, 1931.
108. Domagk, G.: Ein Beitrag zur Chemotherapie der bakteriellen Infektionen, Deutsche med. Wchnschr. **61**:250, 1935.
109. Colebrook, L., and Kenny, M.: Treatment of Human Puerperal Infections and of Experimental Infections in Mice With Prontosil, Lancet **1**:1279 and 1319, 1936.
110. Colebrook, L., and Purdie, A. W.: Treatment of 106 Cases of Puerperal Fever by Sulphonilamide, Lancet **2**:1237 and 1291, 1937.
111. Snodgrass, W. R., and Anderson, T.: Prontosil in the Treatment of Erysipelas—a Controlled Series of 312 Cases, Brit. M. J. **2**:101, 1937.
112. Evans, G. M., and Gaisford, W. F.: Treatment of Pneumonia With 2(p-Aminobenzenesulphonamido) Pyridine, Lancet **2**:14, 1938.
113. Banks, H. S.: Serum and Sulphanilamide in Acute Meningococcal Meningitis, Lancet **2**:7, 1938.
114. Banks, H. S.: Chemotherapy of Meningococcal Meningitis, Lancet **2**:921, 1939.
115. Wagle, P. M., Sokhey, S. H., Dikshit, B. B., and Ganapathy, K.: Chemotherapy in Plague, Indian M. Gaz. **76**:29, 1941.
116. League of Nations Malaria Commission: The Therapeutics of Malaria, Bull. Health Organ., League of Nations **2**:181, 1933.
117. League of Nations Malaria Commission: The Treatment of Malaria, Bull. Health Organ., League of Nations **6**:895, 1937.
118. Fairley, N. H.: Chemotherapeutic Suppression and Prophylaxis in Malaria, Tr. Roy. Soc. Trop. Med. & Hyg. **38**:311, 1945.
119. Fairley, N. H.: Researches on Paludrine in Malaria, Tr. Roy. Soc. Trop. Med. & Hyg. **40**:105, 1946.
120. Black, D. A. K.: Treatment of Burn Shock With Plasma and Serum, Brit. M. J. **2**:693, 1940.
121. Lennox, W. G., and von Storch, T. J. C.: Experience With Ergotamine Tartrate in 120 Patients With Migraine, J.A.M.A. **105**:169, 1935.
122. O'Sullivan, M. E.: Termination of 1,000 Attacks of Migraine With Ergotamine Tartrate, J.A.M.A. **107**:1208, 1936.
123. Wilson, C., Pollock, M. R., and Harris, A. D.: Therapeutic Trial of Methionine in Infective Hepatitis, Brit. M. J. **1**:399, 1945.
124. Higgins, G., O'Brien, J. R. P., Peters, R. A., Stewart, A., and Witts, L. J.: Treatment of Infective Hepatitis With Methionine, Brit. M. J. **1**:401, 1945.
125. Abraham, E. P., Chain, E., Fletcher, C. M., Gardner, A. D., Heatley, N. G., Jennings, M. A., and Florey, H. W.: Further Observations on Penicillin, Lancet **2**:177, 1941.
126. Florey, M. E., and Florey, H. W.: General and Local Administration of Penicillin, Lancet **1**:387, 1943.
127. Pulvertaft, R. J. V.: Local Therapy of War Wounds, Lancet **2**:341, 1943.
128. Porritt, A. E., and Mitchell, G. A. G.: An Investigation Into the Prophylaxis and Treatment of Wound Infection. *In* Penicillin Therapy and Control in 21 Army Group, London, 1945, 21 Army Group, p. 7.
129. Christie, R. V.: Penicillin in Subacute Bacterial Endocarditis, Lancet **1**:369, 1946.
130. Christie, R. V.: Penicillin in Subacute Bacterial Endocarditis, Brit. M. J. **1**:1, 1948.
131. Harrison, S. H., Topley, E., and Lennard-Jones, J.: The Value of Systemic Penicillin in Finger Pulp Infections: A Controlled Trial of 169 Cases, Lancet **1**:425, 1949.
132. Ryle, J. A.: Changing Disciplines, London, 1948, Oxford University Press, p. 104.
133. Anning, S. T.: Calciferol in the Treatment of Chilblains, Lancet **2**:794, 1947.
134. Jackson, D. M., Lowbury, E. J. L., and Topley, E.: Pseudomonas pyocyanea in Burns; Its Role as Pathogen and Value of Local Polymyxin Therapy, Lancet **2**:137, 1951.
135. Jackson, D. M., Lowbury, E. J. L., and Topley, E.: Chemotherapy of Streptococcus pyogenes Infection of Burns, Lancet **2**:705, 1951.

VOL
10
195

# Book Reviews

TUMOURS OF THE BLADDER. Vol. II of NEOPLASTIC DISEASES AT VARIOUS SITES. Edited by David M. Wallace, O.B.E., M.S., F.R.C.S. Edinburgh and London, 1959, E. & S. Livingstone, Ltd. (The Williams & Wilkins Co., Baltimore, exclusive U.S. agents). Pp. 352, indexed. Price $11.50.

OL.
10
959

This scholarly treatise by Dr. D. M. Wallace and his associates is perhaps not a book for medical students. It is rather a book for the student of medicine; for the student of urology. It represents the second volume in the monograph series edited by Dr. D. W. Smithers. Its manner of presentation, scope, and approach are unique and it may best be characterized as a clinical and technical work with numerous statistics. One of its greatest inherent values lies, in this reviewers opinion, in its being a combined effort of the basic and the clinical scientist.

The very method of division of the book implies its purpose and is indicative of the author's different approach to the subject of bladder cancer. The book consists of seven sections, concerned with mortality, industrial risk, biochemistry, examination and classification, pathology, methods of treatment, and analysis of treatment.

The chapters on the industrial worker and industrial chemistry by Goldblatt and the section on etiology and biochemistry by Boyland contain material not available in current texts. Their work represents an outline of the data now available on the relationship of basic biochemistry to the etiology of cancer. In the section on industrial cancer an important aspect of the problem, previously unemphasized, concerns the incidence of tumors in rubber workers.

The sections dealing with clinical material are well presented. There are a few points of possible contention but these are minor.

With regard to the cytology chapter in one of the clinical sections one might at first question the feeling that "the chief and principal use of cytological methods is for the detection of bladder tumors occurring amongst the population exposed to bladder carcinogens," until one examines the usefulness of this technique in their hands. One might, however, more unequivocally disagree with the statement that the Papanicolaou technique was disappointing when used in following recurrent cases with a view to reducing the number of required cystoscopies, since this has been an excellent method, in our hands at the Memorial Center, in following patients with recurrent bladder carcinoma, although it does not take the place of cystoscopy. We are in agreement that the most common mistake in clinical "staging" of bladder cancer has been one of overstaging.

The chapter on mode of death in carcinoma of the urinary bladder is well presented and, in one respect, their approach is more realistic than ours when referring to a "home for the dying." Our terminology for a similar domicile is a "convalescent home."

Chapters on biochemical changes following ureterocolic anastomosis and on radiotherapy of bladder cancer represent excellent review articles. Regarding the chapters on results of treatment, one might question the philosophy that "open or endoscopic diathermy has not been regarded as previous treatment," especially when the authors state, "any patient who received treatment of any sort, even if only one day of external radiotherapy has been included in the treated case."

In summary, this monograph by British clinicians and basic scientists represents an indispensable addition to the library of the urologist dealing with bladder cancer. The fact that many chapters are devoid of a "summary and conclusions" and that many of the tables and charts require detailed study are perhaps an advantage. The book is indeed worth very careful perusal.

Harry Grabstald

HISTORY OF AMERICAN MEDICINE. A SYMPOSIUM. Edited by Felix Marti-Ibañez, M.D. New York, 1959, MD Publications, Inc. Pp. 181, indexed. Price $4.00.

This symposium is made up of contributions from some well-known experts and from some not-so-expert contributors. The result is a book which is uneven and more than a bit repetitious. Nevertheless, there are some interesting chapters, such as the one by Duffy on "Medicine Among the Aboriginal American Indians" and Blair on "The Medicine of Colonial America." Since the foreword states that the symposium has been previously published in the *International Record of Medicine*, perhaps the impoverished reader would be well advised to read the essays in some library and then request reprints of any contributions of particular interest to him. It is unlikely that there will be much of an audience for this little book.

Louis Lasagna

BIOSYNTHESIS OF TERPENES AND STEROLS. Ciba Foundation Symposium, edited by G. E. W. Wolstenholme, O.B.E., M.A., M.B., B.Ch., and Cecilia M. O'Connor, B.Sc. Boston, 1959, Little, Brown & Company. Pp. 311, indexed. Price $8.75.

The Ciba Foundation Symposium on the biosynthesis of terpenes and sterols was held at a most appropriate time since it came during a period of great progress in this field. The contributors to this symposium represent the leading investigators in the field today, and they have done an excellent job of presenting their material.

The most recent "breakthrough" in this problem was the discovery of mevalonic acid by Dr. K. Folkers and his co-workers in 1956, and this topic was presented by Dr. Folkers. Dr. K. Bloch presented his elegant work on the further metabolism of mevalonic acid and its conversion by a yeast enzyme system to squalene via phosphorylated intermediates. Some of the other topics discussed included the biosynthesis of branched-chain acids, the biosynthesis of squalene and cholesterol in liver enzyme systems, alterations of cholesterol biosynthesis in liver cell fractions from rats in various experimental conditions, the formation and metabolism of bile acids, and the biosynthesis of terpenes and carotenoids in micro-organisms, fungi, and higher plants.

This record of the Symposium contains a number of experimental results which had not previously been published, and it should be of great interest to those working in the field and also to those engaged in other aspects of terpene and sterol chemistry and biochemistry. In view of the possible importance of one sterol, cholesterol, in aberrant lipid metabolism and atherosclerosis, this subject is also of great interest to workers in other disciplines, and they should find this book a valuable reference.

James W. Ogilvie

BLOOD GROUPS: *British Medical Bulletin*, Volume 15, Number 2, May, 1959, pages 89-174. Price $3.25. (Published by the Medical Department, The British Council, 65 Davies Street, London, W. 1.)

It is entirely fitting that a review of the subject of blood groups should appear in the *British Medical Bulletin* and should be written by a group of British investigators. The wide range of interest of this subject is clearly demonstrated by the backgrounds of the authors. Dr. Mourant's introduction and his paper on the relation of blood groups and anthropology are important contributions to this Symposium. The paper on inheritance of blood groups is particularly lucid and the other papers in the Symposium are all well written and informative. An additional extremely valuable feature of this Symposium is the selected list of references following each paper. The entire review is of the same high quality which was seen in the preceding issue, "Haematology."

Julius Krevans

NEUROPHARMACOLOGY. Transactions of the Fourth Conference, Sept. 25-27, 1957. New York, 1959, Josiah Macy, Jr. Foundation. Pp. 285, indexed. Price $5.00.

For half of this book, I had the unpleasant feeling that the Macy Foundation had, in a moment of lunacy, restarted the old "Problems of Consciousness" series. Dr. Abramson starts things off

with some more unexciting data on the effects of drugs on young Siamese fighting fish. (They're so young that they apparently have no desire to fight, which is more than you can say for the participants at the conference.) Then Heath and his crew from Tulane present, at tedious length, their "taraxein" and EEG studies on schizophrenics. One can almost sense in the transcript the hostility and suspicion of Heath's audience, and his defensive and humorless reaction. Although there is plenty of give-and-take in this section, one wonders at the silence of some of the audience (Molly Brazier, for example, who is a remarkably competent and experienced electroencephalographer, says almost nothing about Heath's deep electrode data).

Fortunately, John Lilly follows with a concise, interesting, and delightful presentation of "pain" and "pleasure" areas in the monkey brain. The book ends on a downhill note, with Gordon Alles' rambling, unclear presentation of some chemical and biologic data on some phenethylamines. Once again, Heath gets into the act here (along with Abood) to provoke a wierd Alice-in-Wonderland discussion of chemical tests for schizophrenia. To cap it all off, the editor has allowed Grenell (who said almost nothing at the conference) to add 5½ pages of small type, as an "afterthought." This would not be so bad were it not for the fact that the "afterthought" is essentially a set of musings (about chemical structure and drug action) which seem rather pointless at this stage of our knowledge.

Question: Would the Macy Foundation not do more good with its money if they skipped conferences like this one and supported research in the area with the funds thus saved? It would also cut down on the time wasted by some of the younger participants, who could be home "minding the store."

Louis Lasagna

MIRAGE OF HEALTH. By René Dubos. World Perspectives, Vol. XXII. New York, 1959, Harper and Brothers. Pp. 236, unindexed. Price $4.00.

If one were asked to nominate the ten most provocative living scientists, the list would almost certainly include René Dubos. His extraordinary range of interests and felicitous prose have made him equally well-known among laymen and scientists. The present book is an interesting continuation of some of Dr. Dubos' past writings, and can be highly recommended.

Here, for example, are some samples of Dubosian prose:

"Is it not a delusion to proclaim the present state of health as the best in the history of the world, at a time when increasing numbers of persons in our society depend on drugs and on doctors for meeting the ordinary problems of everyday life?"

"Someday, no doubt, love philters will be sold over the counters, drugs more dependable than whiskey will help timid men to behave for a while like heroes, and other drugs may be used to direct life toward goals that will determine the future trend of human evolution." (One is tempted to ask whether anything will ever replace alcohol, the thinking man's philter.)

"Abstract concepts such as resistance and susceptibility of the host, virulence, attenuation, invasiveness, and toxicity of the parasite have all been invoked as if they corresponded to real entities."

"Thus, the mechanisms which increase permeability of the gut might be regarded as the real cause of [cholera] since they account for its symptomatology and since treatment of the effects usually results in cure."

"Like other forms of disease, most cancers are probably the result of a constellation of circumstances. This complexity appears disheartening to the investigator, but it is also a source of hope, since the multiplicity of links in the chain of causation increases the chance of finding a weak part in the chain and thus of discovering a method of control."

"In our societies, the school of public health always plays second fiddle to the school of medicine."

"It is a remarkable fact that the greatest strides in health improvement have been achieved in the field of diseases that responded to social and economic reforms after industrialization. The nutritional deficiencies that were so frequent in the nineteenth century have all but disappeared in the Western world, not through the administration of pure vitamins but as a result of over-all better nutrition. The great microbial epidemics were brought under control not by treatment

with drugs but largely by sanitation and by the general raising of living standards. In contrast, the cancers, the vascular disorders, the mental diseases, which were not affected by the sanitary movement, have remained great health problems and their solution is not yet in sight. It is legitimate to hope, of course, that vigorous research will yield drugs for the relief of patients suffering from these diseases, but it can be predicted that drug treatment will not provide the real solution to the problem."

". . . the empirical hit-or-miss attack dignified in scientific circles by the expression 'screening program.' "

But it would be incorrect to give the impression that Dr. Dubos hits the bull's-eye every time. There are a number of opinions expressed that seem less soundly based on fact. For example, I think it unlikely that country clubs, ski lodges, beach bungalows, clambakes, and barbecues represent attempts by twentieth century man to recapture the spirit of Rousseau, nor do I believe that hors d'oeuvres served at cocktail parties reflect "biological wisdom." The neo-Darwinism of the author seems (at least to one nonexpert) remarkably naïve. Is it indeed likely that for *man* "biological fitness achieved through evolutionary adaptation . . . [is] the most dependable touchstone of permanent success . . ."?

My experience with dogs and cats may have been idiosyncratic, but I could hardly agree that "even the most pampered cat or dog will happily eat day in and day out a standardized diet canned in Chicago . . ." The wild Norway rat does indeed have ". . . an extraordinary ability to resist all sorts of stresses" but what of his *increased* sensitivity to certain stresses, such as handling? And what does this sentence mean? "In a wise society leisure and holidays—instead of becoming stereotyped as they presently are—should play a role similar to that of national parks and wildlife reservations, where plants and animals retain some chance to practice the mechanisms which have permitted evolutionary adaptation."

Finally, it appeared to this reviewer that the most fascinating and delightful parts of the book were the middle sections, where Dr. Dubos was covering those areas in which he has been most interested and in which he is so extraordinarily competent. The early and late chapters read less smoothly and seem in some ways tacked on to provide a beginning and end. It is quite likely, however, that other readers will have exactly the opposite reaction, so for Heaven's sake don't let the foregoing criticisms keep you from reading this book. To paraphrase Alexander King: If you can't enjoy this book, man, you'd better consult your minister, priest, rabbi, or psychoanalyst, because you're *sick!*

Louis Lasagna

# Announcements

## UROLOGY AWARD

The American Urological Association offers an annual award of $1,000 (first prize $500, second prize $300, and third prize $200) for essays on the result of some clinical or laboratory research in Urology. Competition is limited to Urologists who have been graduated not more than 10 years, and to hospital interns and residents doing research work in Urology.

The first prize essay will appear on the program of the forthcoming meeting of the American Urological Association, to be held at the Palmer House, Chicago, Ill., May 16-19, 1960.

For full particulars, write the Executive Secretary, William P. Didusch, 1120 North Charles St., Baltimore, Md. Essays must be in his hands before Dec. 1, 1959.

VOL.
10
959

## POSTGRADUATE CEREBRAL PALSY COURSES—FALL, 1959

Columbia University, College of Physicians and Surgeons, New York, present their fall schedule for postgraduate cerebral palsy courses. These courses are designed (1) to acquaint potential leaders in this field with basic knowledge concerning cerebral palsy; (2) to review the present status of practical therapy, including the medical, social, educational, and psychological aspects; and (3) to point out the controversial features of our present understanding.

The first 3 weeks are devoted to these lectures and demonstrations:

Medical management (53 hours)
- *Astereognosis
- *Bracing
- *Clinical research
- Clinical signs
- *Evaluation clinics
- *Follow-up case studies
- *Hearing and language
- *Pediatrics
- *Present trends
- *Seizures
- *Surgery
- *Therapy aims and techniques
- *Vision

Basic science (22 hours)
- Kinesiology
- Neuroanatomy
- *Neurophysiology

Psychosocial aspects (13 hours)
- *Educational methods
- *Psychiatric problems
- *Parent counseling
- *Psychometrics
- *Vocational problems

Organization (2 hours)
- Community resources
- *Public health

The last 6 weeks of therapists' and nurses' course are spent as follows:

22 days in clinical experience in all types of facilities within the New York area: diagnostic, inpatient, outpatient, rehabilitation (physical and vocational), preschool to adult recreation, school, custodial

7 days of seminars for discussion of reading and clinical experience, covering: handedness, multiple handicaps, parent-therapist relations, records, recreation, various therapeutic approaches, and requested topics.

Informal questioning of lecturers is encouraged. Through "give and take" discussions, in class and in clinics, students enjoy opportunities to clarify their ideas, to learn from the experience of others, and to evaluate treatment techniques of various therapeutic approaches.

*Physicians' Course: 3 weeks full time. Oct. 5-23, 1959. Fee $100. Includes lectures and demonstrations plus 5 days in cerebral palsy facilities with opportunity for discussion with the physician in charge. Starred (*) sessions are included in physicians' course.

Therapists' and Nurses' Course: 9 weeks full time. Oct. 5-Dec. 4, 1959. 8 points credit. Fee $210. For occupational, physical, and speech therapists and nurses with proper qualifications and experience.

A list of scholarship sources will be sent on request.

# COURSE IN REHABILITATION CARE OF THE CHRONICALLY ILL PATIENT

A one-week course for physicians, devoted to the rehabilitation care of the chronically ill patient, will be held Nov. 16-20, 1959, under the auspices of the Department of Physical Medicine and Rehabilitation, New York Medical College—Metropolitan Hospital Center. The course will offer a review of the principles and techniques in the medical care of the chronically ill to meet the needs of the clinician, medical administrator, and Public Health physician. Course content will include: physiology and pathology of chronic diseases, nutrition and dental care, management of bedridden and incontinent patients, home care programming, community needs and resources, Public Health aspects, self-care activities, prosthetic devices, and psychological and social aspects.

The tuition fee is $100.00. Traineeships for tuition, maintenance, and travel are available through funds provided by the U. S. Office of Vocational Rehabilitation. Applications for the course and traineeships can be obtained directly from Dr. Jerome S. Tobis, Chairman, Department of Physical Medicine and Rehabilitation, New York Medical College, 1 East 105th St., New York 29, N. Y.